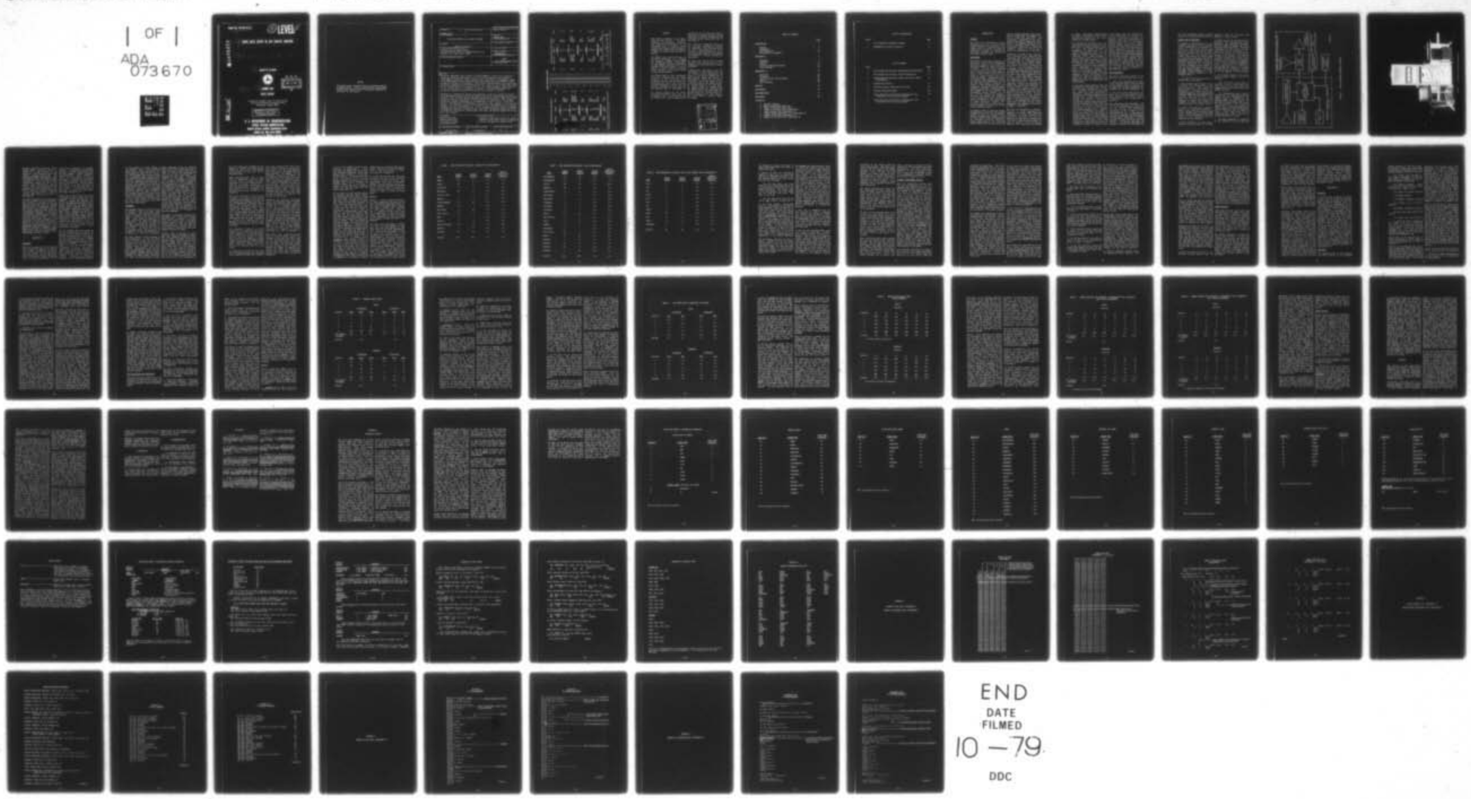
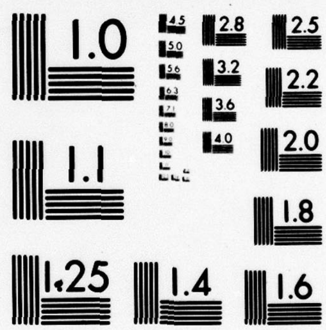
AD-A073 670 NATIONAL AVIATION FACILITIES EXPERIMENTAL CENTER ATL--ETC F/G 17/2
VOICE DATA ENTRY IN AIR TRAFFIC CONTROL.(U)
AUG 79 D W CONNOLLY
UNCLASSIFIED FAA-NA-79-20 NL

1 OF 1
ADA
073670



END
DATE
FILMED
10-79
DDC



MICROCOPY RESOLUTION TEST CHART

Report No. FAA-NA-79-20 (14)

(11) LEVEL II

AD A073670

(6) VOICE DATA ENTRY IN AIR TRAFFIC CONTROL.

(9) Rept. for May 75 - Dec 78,

(10) Donald W. Connolly

(12) 77p.

(11) AUGUST 1979

NAFEC REPORT

Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

U. S. DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
National Aviation Facilities Experimental Center
Atlantic City, New Jersey 08405

NOTICE

The United States Government does not endorse products or manufacturers. Trade or manufacturer's names appear herein solely because they are considered essential to the object of this report.

Technical Report Documentation Page

1. Report No. FAA-NA-79-20	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle VOICE DATA ENTRY IN AIR TRAFFIC CONTROL		5. Report Date AUGUST 1979 6. Performing Organization Code
7. Author(s) Donald W. Connolly		8. Performing Organization Report No.
9. Performing Organization Name and Address Federal Aviation Administration National Aviation Facilities Experimental Center Atlantic City, New Jersey 08405		10. Work Unit No. (TRAIS) 11. Contract or Grant No. ~~219-151-100~~
12. Sponsoring Agency Name and Address		13. Type of Report and Period Covered NAFEC May 1975-December 1978 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Two major experiments and a number of subsidiary pilot studies were conducted to assess the potential operational utility of state-of-the-art word recognition technology in air traffic control applications. Experiment I, employing 12 operators or "talkers", secured baseline data representing the inherent "best case" recognition accuracy of the system. Three of the subvocabularies of an operational data entry language were tested exhaustively to a total of over 46,000 spoken words. On the average, across all speakers and all three subvocabularies, only 1 percent of the words spoken were erroneously recognized. Subsequent "tuning" of the recognition algorithm reduced the error rate to less than 0.4 percent.

Experiment II compared the quality and efficiency of the voice system versus the existing keyboard method of entering complete operational messages. Five operators entered a total of 6,000 complete messages using the two methods. The voice system produced 64 percent fewer errors of all kinds than the keyboard in the messages entered but showed no advantage (or disadvantage) in rate of entry. Voice entry, however, did demonstrate an advantage in the overall time consumed by the data entry process - less time was required to "translate" entry requirements into the spoken language than into the keyboard code. Other influential factors were also examined, including the effects of microphone type, individual differences between operators, the stability of operators' "word-prints" over time, and operator adaptation to the voice entry system.

17. Key Words Speech Processing Computer Data Entry Air Traffic Systems Human Engineering		18. Distribution Statement Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 75	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

METRIC CONVERSION FACTORS

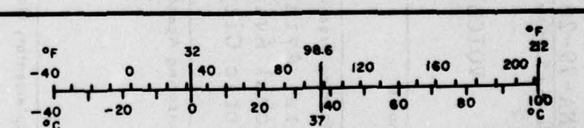
Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
	LENGTH			
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
	AREA			
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
	MASS (weight)			
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
	VOLUME			
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
	TEMPERATURE (exact)			
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in ≡ 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, *Units of Weights and Measures*, Price $2.25, SD Catalog No. C13.10:286.

Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
	LENGTH			
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
	AREA			
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
	MASS (weight)			
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
	VOLUME			
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
	TEMPERATURE (exact)			
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



PREFACE

Many people cooperated in the completion of this project. The whole community of government, academic, and industrial researchers in the expanding field of computer speech processing provided continuing encouragement and consultation throughout. A number of people deserve special thanks.

Mr. Robert Cox, Chief Engineer of Threshold Technology Inc., designed the syntactic software which formed the operational basis of experiment II, the improved training system which evolved from experiment I, and a number of other creative programs which greatly facilitated data collection and analysis.

Mr. Wilfred Rawl of Data Transformations Inc. converted all the system software to the real-time disk-operating system and developed the software for the voice storage/retrieval/feedback system and the powerful programs for data reduction and analysis used in experiment II.

Dr. Robert Breaux of the Naval Training Equipment Center and his colleagues, Mr. Ira Goldstein and Mr. James Duva, provided immeasurable assistance all the way from aid in securing the original hardware through many hours of patient consultation and technical advice during the life of the project.

Mrs. Florence Champion served as the principal experimenter throughout experiment I and the long period of pilot experimentation which followed, contributing much in the line of insight toward the development of methods of system reliability improvement.

Mr. Paul Quick and Mr. Scott Stemple designed and fabricated the very original and functional voice response subsystem.

Perhaps most important of all who contributed to this project were the air traffic control specialists and others of the staff of NAFEC and the FAA Eastern Region who served as experimental operators--willingly, cheerfully, and in the spirit of true scientific inquiry: Mr. Edward Alzner, Mr. Rudolph Antonio, Mr. John Badzo, Mr. Robert Clark, Mr. Anthony Coia, Mr. Edward Ezekiel, Ms. Arden Lansing, Mr. Robert Lucas, Mr. David Mark, Mr. Charles McGee, Mr. Charles Moore, Mr. Melvin Rosenbaum, Ms. Adele Stahley, Mr. Mark Taylor, and Mr. Wayne Willoughby.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

INTRODUCTION

PURPOSE.

The goal of the experimental studies reported here was to assess the state-of-the-art in recognition of the spoken word by means of computer technology in order to evaluate its potential usefulness in operational air traffic control (ATC).

BACKGROUND.

The initial operational capability of automation in air traffic management was never expected to provide the absolute ultimate in either efficiency or safety. It was, rather, intended to set in place the foundations on which continuing evolutionary improvements in both areas could be built while at least keeping pace with predicted demand. Substantial effort is currently being expended to improve the quality and completeness of the raw data base which is essential in order to reap optimum benefit from automation, as witness the development programs in beaconry, communications, and navigation. For the present and foreseeable future, however, one of the critical sources of complete, accurate, and timely data regarding the instantaneous and projected traffic situation is the large number of human operators (traffic controllers and their assistants), several thousand of whom are on duty at any instant in time.

At present, there is only one channel through which controllers can transmit essential facts to the automation system: through their fingers. Many of these critical items of information are available from the controller alone, being based on (or resulting from) decisions he has made on the one hand or having originally been transmitted to him (as, for example, by pilots under his control). The keyboard "language" that must presently be used to communicate these data to the computer system is artificial, encoded, almost absolutely inflexible, difficult to learn and remember, subject to error, and a source of distraction to the user.

Automation in traffic control systems, even initially, has improved the quantities and qualities of information available to controllers while relieving them (to a limited degree) of some of their former mental, vocal, and manual activity (references 1 and 2). It is no longer necessary to remember target identities, nor is frequent radio communication necessary for acquiring or reacquiring identity, altitude, and speed information from pilots. These advantages, however, are secured at least in part through the imposition of new or altered tasks upon the controller. He must now manipulate switches and keyboards--to modify the content of his display, to execute certain control actions, and to update the computer store which provides him with the improved information in the first place. The data entry workload has, in fact, necessitated increased sector staffing in a number of enroute control facilities. Thus, a new language has been introduced into the world of air traffic management--the language of data entry messages to computers.

The fact of the matter is, however, that all of these "messages" are composed in natural human language, formulated in words, phrases, and sentences and many (if not the majority) of them must be communicated from man to man by human speech as well as entered into the digital computer files upon which the system

is based. The human language which is used, furthermore, is a much restricted subset of the total repertoire of human speech.

The substantive vocabulary for any specific ATC operator position is of the order of three hundred "words" or less. The number of kinds of "messages" or "sentences" is, of course, substantially less than the number of "words" (reference 3). The structure of ATC verbal messages is also rather rigid. All of these factors--the small vocabulary, limited message set, and strict syntax--tend to reduce the problem of speech recognition to one of more manageable size. Furthermore, it is not necessary to be able to interpret the speech of any speaker whatever, but only of a limited number of known speakers; to wit, the specific controllers at specific positions at a given time. Nor is recognition of the speaker required, for essentially the same reasons.

It has been widely observed that the technology of isolated word recognition is "here" (reference 4, 5, and 6). "Isolated" in this context means only that the word must have a definable, detectable beginning time and end time. A "word" may be multisyllabic or, indeed, a rather long phrase so long as it is uttered continuously without detectable stops. Current techniques in this class are capable of "word" recognition accuracies of well over 90 percent with known speakers (i.e., speakers who have "pretrained" the device to their own vocal idiosyncrasies by speaking each "word" in the vocabulary several times) and achieving moderate sized vocabularies (e.g., 10 to about 50 words). Accuracies of 98 to 100 percent have been obtained where the vocabulary consists of digits alone.

A substantial part of ground-to-air voice communications (about 20 percent) consists precisely of numbers, while the "vocabulary" of keyboard data entry in the model A-3d-2 (Model A) enroute system consists almost entirely of numbers and letters. Thus, the application of voice recognition in air traffic control does not necessarily require interpretation of discursive conversation or much (if any) "understanding" of "continuous speech" in an unlimited (or even very large) language. While many opinions have been advanced regarding the applicability of voice data entry in ATC systems, the fact of the matter is that the question has not yet been systematically, experimentally tested.

TEST OBJECTIVES.

The basic questions asked of the studies reported here were two:

1. Given the vocabulary of an operational ATC data entry function, what is the highest order of accuracy (or "reliability") of word recognition obtainable with current technology, and

2. How does voice data entry compare with existing keyboard entry with regard to accuracy, speed, learnability, and acceptance by operators?

Two experiments were performed. The first was designed to determine (a) the inherent word-recognition accuracy of the best obtainable technology using a number of the subvocabularies of a representative data entry language from the National Airspace System (NAS) and (b) methods of improving recognition accuracy wherever it might be found less than perfect. The second experiment was designed to compare the performance of

the word recognition method of data entry to the existing keyboard method.

DESCRIPTION OF EQUIPMENT.

A schematic representation of the equipment used may be found in figure 1 and a photograph of the assembly in figure 2. The basic device used was a Voice Input Processor, model VIP-100, manufactured by Threshold Technology, Inc., of Delran, New Jersey. This device was chosen based on a National Aviation Facilities Experimental Center (NAFEC) survey of available systems and on the basis of surveys performed for the Naval Training Equipment Center (NTEC) as well as the experience of NTEC with the same model equipment in studies performed at NTEC (reference 7). This equipment functions generally in the following manner:

1. A single, univocal utterance is spoken into the microphone.

2. The waveform audio energy of the utterance is transmitted to the audio digitizer. The digitizer incorporates 32 audio filters or "features," 16 of which are of the frequency/amplitude type spanning the frequency range from approximately 250 through 5250 hertz (Hz). The other 16 filters are specially designed to detect the presence of composite or unique sounds which are characteristic of human speech.

3. Every 2 milliseconds (approximately) the digitizer delivers a 32-bit (one per filter) digital image of the immediately preceding audio signal to the minicomputer, from the onset of the utterance to the cessation of the utterance.

4. The software in the minicomputer stores and saves all of these "samples" from the digitizer and maintains a count of them.

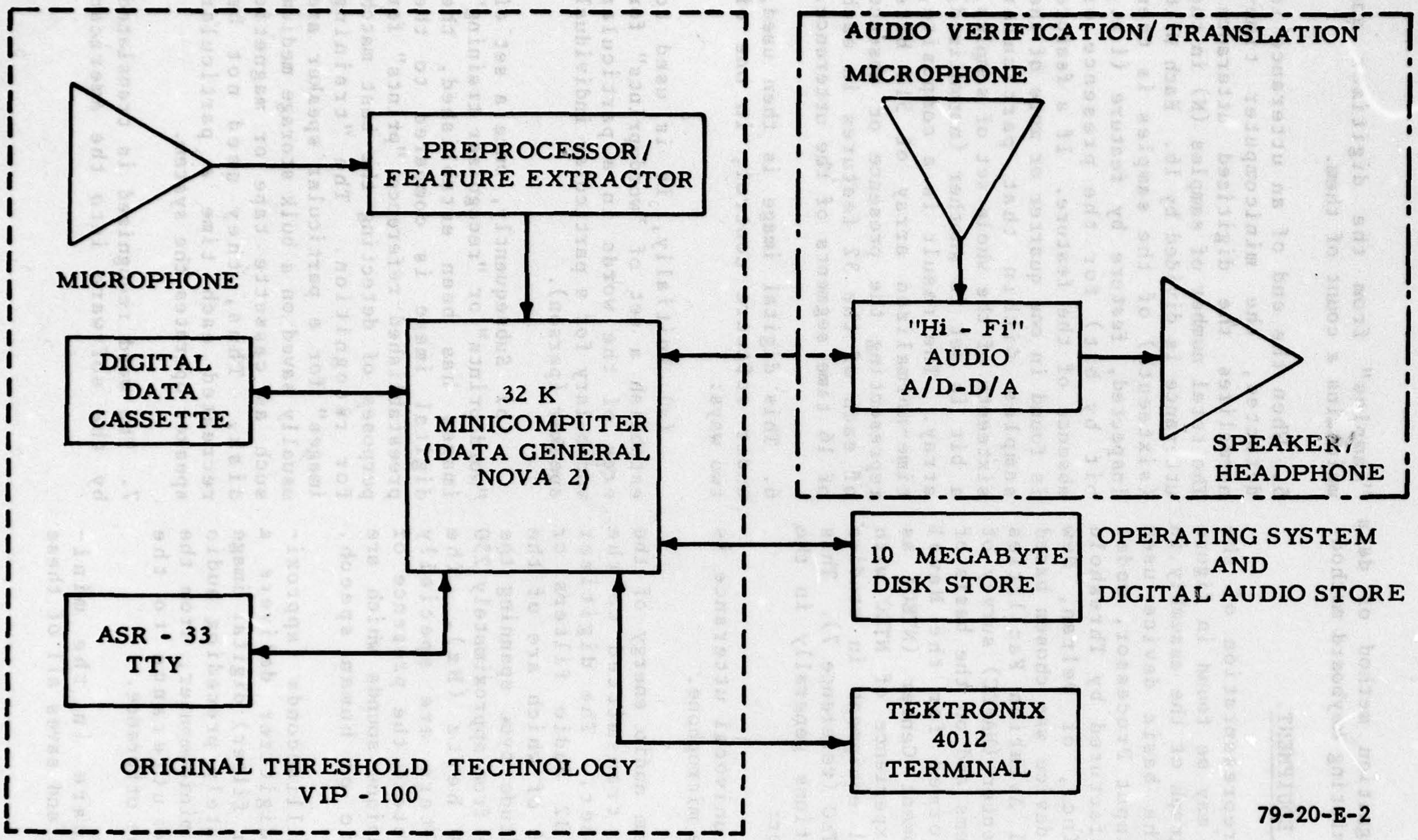
5. When the end of an utterance is detected, the minicomputer time-normalizes the digitized utterance. The total number of samples (N) in the utterance is divided by 16. Each N/16 (sixteenth) of the samples is then inspected, feature by feature (i.e., bit by bit) for the presence or absence of the feature. If a feature is found in one quarter or more of the samples within that particular sixteenth of the whole set of samples, a bit is set in another (normalized) array. The result is a composite, time-normalized array of 512 bits representing the presence or absence of each of the 32 features in each of 16 time segments of the utterance.

6. This digital image is then used, under software control, in one of two ways:

(a) Initially, it is used to establish a set of "word-prints" for each of the words in a particular vocabulary for a particular individual speaker (person).

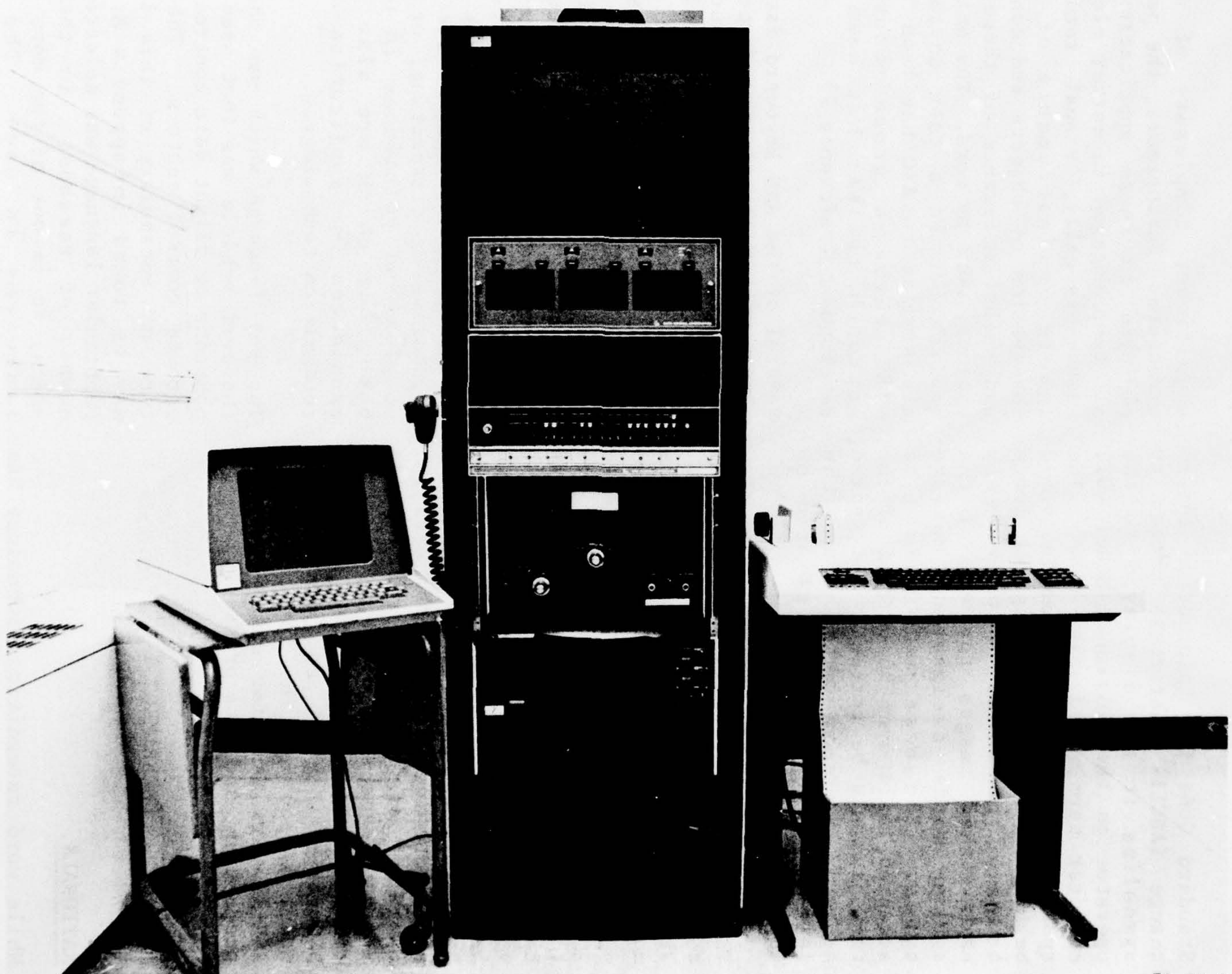
(b) Subsequently, once a set of "word-prints" or "recognizer training" images has been established, the digital image is compared to the preestablished reference "prints" for purposes of detecting the best match for recognition. The "training images" for a particular speaker are usually saved on a bulk storage medium such as cassette tape or magnetic disk. Thus, they need not be recreated each time a particular speaker operates the system.

7. The word recognized is translated by the software into the American



79-20-E-2

FIGURE 1. VOICE LABORATORY SCHEMATIC DIAGRAM



79-20-2

FIGURE 2. FAA/NAFEC VOICE ENTRY LABORATORY

Standard Code for Information Interchange (ASCII) characters chosen to symbolize it and displayed to the operator on the Tektronix Model 4012 computer terminal display tube. The spoken word "Amend" for example, would be digitized, and the digital image compared to a set of previously established images for all of the spoken words designating message types. The best match would be accepted (well over 96 percent of the time it would be correctly recognized as will be seen below), and the letters "AM" would be printed on the display.

The other equipment in the laboratory system served various purposes. The cassette transports and the magnetic disk were used for storage and retrieval of programs, experimental data, and "training" data for the various operators who served as experimental subjects. The audio verification subsystem was designed and constructed at NAFEC. It was briefly evaluated for use as a substitute for visual feedback to the operator of the word recognized. More will be said of this below in connection with experiment II. The DECwriter (Digital Equipment Corp., model LA-36) and the Teletype model ASR-33 were used for system control, programing, and data printout.

EXPERIMENT I

RATIONALE.

While word recognition technology has been successfully applied in a significant number of commercial operations such as package routing and manufacturing quality control inspection, none of these have involved languages with very large vocabularies or any significant complexity of message structure. Furthermore, the personnel employed in these applications have all been engaged in rather elementary tasks--basically, visual reading of labels or instruments or visual observation of objects and conditions and verbal utterance of these observations, word by word. The only known application in a task anything at all similar to ATC has been that of NTEC where a ground-controlled approach trainer has been under development (reference 7).

Several of the NAS keyboard data entry "languages" were tabulated and analyzed. There are two such languages in regular and extensive use in the semiautomated enroute traffic control centers of the agency which produce daily hundreds of thousands of messages requiring millions of keystrikes. There are a number of other entry languages in the system (e.g., control tower cab, terminal radar control facility, flight service station, etc.) which are either not as burdensome or distracting, or not as complex and voluminous in use, or both, but which are also likely candidates for application of word recognition technology.

The key language which was chosen as the test vehicle was that used by the nonradar or flight data controllers in enroute control centers. The structure and vocabulary of this language may be found in appendix A. This particular language was selected for a number of reasons. In the first place, it is one of the more complex languages in use. The total repertoire of possible messages is larger than that of any of the other key languages used by personnel engaged in the active control of traffic. Finally, the key-entry workload at this operational position

is the largest, in total volume, in the system. Thus, a very difficult application was undertaken for investigation right at the outset. The theory behind this choice was that (a) it appeared highly likely, given the state of the word recognition art, that this application would be practical and cost/beneficial and that, a fortiori, less complex, less difficult applications would yield to the same approach with zero or minimum additional research and development effort or that (b) many or most of the relevant questions for the lesser applications would be answered in the course of attacking the greater, even if the present state of technology did not prove practical for this particular application.

PROCEDURE.

The language chosen for test was found to include a total of 24 basic types of message. (An additional seven types of message covering "conflict alert" entries have since been added. This, based on experience to date, should not cause any special difficulty.) Of these, 15 types of messages encompass 96 percent of all messages actually entered in operation. In addition, these 15 message types (see appendix A) include all of those occurring with a frequency of 1 in 100 or greater.

The first element of every message is the message type. It was also found that, in most cases, the type of message must be followed by the identity of the flight data file (flight plan) to which the entry applies. Furthermore, of the four means of identifying a flight, the one most commonly employed was the three-decimal-digit computer identity number assigned to every flight (reference 7). Thus, the second element of most spoken messsages could be assembled from a word list consisting only of digits plus two or three control words (such as "erase" for restarting the whole entry and "backspace" for changing the last digit). The second element of some types of message (e.g., weather information retrieval) and the third or fourth element of other messages (e.g., early handoff to a terminal; hold message) is a location identifier or geographic "fix." The keyboard codes for these place names are not always mnemonic (e.g., Benton is coded 7QB), but the place names themselves are easily spoken. No attempt was made to survey all possible fix-names; however, those tested here included, for one sector in the New York air route traffic control center (ARTCC), all VOR's, all intersections, and all terminals; in short, all the fixes normally required at the position as elements of key-entry messages.

Two types of message (flight plan amendment and correction) require identification or naming of a flight plan data field (e.g., assigned altitude, speed). Eight of these data fields account for the vast majority of modifications entered, and the field content or substantive data most commonly consists of digits.

Certain types of entries or, more precisely, parts of messages currently made with keyboards basically exist only in coded, nonverbal or partially nonverbal form. Consider the aircraft identity N1009Y (tail number). The most convenient way to make such an entry might still be via keyboard. However, an "all purpose" subvocabulary consisting of all of the digits plus the phonetic alphabet (which is part of the linguistic stock-in-trade of the air traffic controller) was made a part of the total vocabulary of

the voice data entry language for the purpose of making the comparatively fewer and rarer entries not already encompassed by the word lists described above.

These subvocabularies, plus a short list of commercial aircraft types and the list of relevant avionics equipments (or type "Qualifiers"), make up the whole vocabulary as currently constituted. The vocabulary and syntax of the language, as previously noted, are included here in appendix A.

The first experiment conducted was intended to establish the basic recognition performance of the VIP-100 word recognition package with three of the subvocabularies discussed above; namely, the 15 message types, the 21 fix names, and the 10 digits (plus "erase" and "backspace") list. Each of the lists, separately, was expanded into a pseudorandom assembly in which each member of the list appeared 10 times. The list used for the digits subvocabulary may be found in appendix B. Thus, the "reading list" for message types was 150 "words" long; for digits, 120 words; and for fixes, 210 words.

Each speaker "trained" the word recognizer by speaking each "word" (some, as may be seen the appendix B, were composites or phrases spoken without internal pauses) 10 times. This resulted in composite digital images of the way the speaker speaks each of that particular list of words. These reference images were then written on cassette tape for later reuse.

It should be pointed out here that this "training" process was conducted by the same-word-repetition procedure that was essentially built into the system program as delivered. Other users of, and experimenters with, this type of equipment use this same procedure. For example, for the vocabulary consisting of digits, the new operator first speaks the word "zero" 10 times in succession with a brief pause (approximately one-tenth of a second) between successive enunciations. Then the word "one" would be spoken 10 times and so forth through the word "nine." This is an important point to note in the light of discoveries which were later made during attempts to improve the accuracy of the system.

Following the initial "training" session, each speaker read the pseudo-random list described above (now for recognition) in 10 separate sessions, in the case of message types and fixes, 5 sessions for the digits list. Data were automatically collected during each test session on the number of times each word was correctly recognized by the computer, the number of times incorrectly recognized, the average closeness of match between the spoken entry and the best and second-best choices among the reference images (i.e., the training images), and the duration of the spoken expression. Samples of raw and processed data are in appendix A.

Each subject, over a period of several days to several weeks, spoke (for recognition testing) each word in each of the subvocabularies 100 times for the types and fixes and 50 times for the digits. The principal purpose of testing digits was to ascertain whether the sample of speakers produced the order of recognition accuracy for digits which is commonly found using this word recognition equipment.

A total of 12 speakers served as test subjects for experiment I. Nine were male journeymen ATC specialists with extensive experience in the NAS Enroute Test Facility. Three were noncontrollers, two female and one male. One group of 11 of these speakers served as subjects for the message types (9 male, 2 female) and another group of 11 (10 male, 1 female) from the same pool of speakers served for the fix names and the digits subvocabularies.

In the matter of user familiarization and operator training, several important observations were made. During the test series for each speaker with each word list, recognition accuracy and "rejection" data were processed at least after every second session. As a rule, in the event that any individual word was either erroneously recognized two or more times or rejected as unrecognizable two or more times, a new set of "training" data was made for that word (and for the word with which it was confused if the confusion was consistently between the same two words). Thus, as recognition testing proceeded, the quality of the reference images or "training data" for some of the words in each list for some of the speakers was progressively refined. This does not mean that a great deal of retraining was done. A number of the speakers never needed to "retrain" any of the words in any of the lists at all. For example, on the <u>average</u>, each speaker needed to retrain one word one time for the list of fixes. Some speakers needed to retrain more words than others, and some of the words and word pairs were more troublesome than others; for example, the fixes Milton and Benton in the list of fix-names. Attempts by some speakers to adopt an extraordinary (for them) pronunciation or emphasis in an attempt to improve recognition of a word had the reverse effect. Habitual or "natural" expression of the utterances is vital to accuracy of recognition.

It should be pointed out that the operators did not receive feedback of results during testing. The experimenter could see the feedback display but the operator could not. The only indication operators received about results came to them very indirectly when they were asked to retrain a word or words as noted above.

<u>RESULTS</u>.

Tables 1, 2, and 3 contain the results for word recognition accuracy of the basic Threshold Technology system for the three subvocabularies tested. Each entry in tables 1 and 2 is based on 1,100 voice entries. Each of the 11 speakers spoke each word for recognition by the system 100 times. Each entry in table 3 is based on 550 repetitions of each word--each word spoken 50 times by each of 11 speakers.

The basic data represented in tables 1, 2, and 3 are the numbers of times each word was misrecognized as some other word in the same subvocabulary (i.e., errors) and the number of times the word was rejected (i.e., not accepted as any of the words in the subvocabulary). There are, obviously, several ways that "accuracy" could be defined in this situation. An error (misrecognition) by a voice entry system is certainly undesirable, indeed totally undesirable. Rejects, or "refusals" to recognize the utterance at all cannot cause direct harm. If, however, the rejection rate is very high (for example, one out of two utterances rejected) even if there are no errors at all it would require

TABLE 1. WORD RECOGNITION ACCURACY: MESSAGE-TYPE SUBVOCABULARY

WORD	NUMBER ERRORS	NUMBER REJECTS	PERCENT ERRORS	PERCENT ERRORS PLUS REJECTS
Amend	25	27	2.3	4.7
Cancel	62	21	5.6	7.5
Correction	3	2	0.3	0.4
Departure	20	75	1.8	8.6
Discrete Code	1	7	0.1	0.7
Readout	1	8	0.1	0.8
Accept Handoff	27	48	2.5	6.8
Handoff	9	3	0.8	1.1
Drop Track	4	20	0.4	2.2
Print Strip	13	8	1.2	1.9
Hold	2	7	0.2	0.8
Release	0	0	0.0	0.0
Report Altitude	21	45	1.9	6.0
Weather	7	11	0.6	1.0
Transmit	32	10	2.9	3.8
Overall:	227	295	1.4	3.2

TABLE 2. WORD RECOGNITION ACCURACY: FIXES SUBVOCABULARY

WORD	NUMBER ERRORS	NUMBER REJECTS	PERCENT ERRORS	PERCENT ERRORS PLUS REJECTS
Williamsport	35	8	3.2	3.9
Selingsgrove	12	9	1.1	1.9
Milton	39	93	3.5	12.0
Hazelton	28	27	2.5	5.0
Wilkes-Barre	3	6	0.3	0.8
East Texas	1	8	0.1	0.8
Lake Henry	1	6	0.1	0.6
Tobyhanna	6	5	0.5	1.0
Allentown	2	5	0.2	0.6
Stillwater	2	3	0.2	0.5
Benton	15	43	1.4	5.3
Sweet Valley	1	0	0.1	0.1
Lopez	2	0	0.2	0.2
Snyders	3	1	0.3	0.4
Slatington	7	2	0.6	0.8
White Haven	30	4	2.7	3.1
Resort	8	23	0.7	2.8
Pennwell	3	13	0.3	1.5
Huguenot	17	15	1.5	2.9
Solberg	10	2	0.9	1.1
Freeland	9	17	0.8	2.4
Overall:	224	290	1.0	2.2

TABLE 3. WORD RECOGNITION ACCURACY: DIGITS AND CONTROL WORDS SUBVOCABULARY

WORD	NUMBER ERRORS	NUMBER REJECTS	PERCENT ERRORS	PERCENT ERRORS PLUS REJECTS
Zero	1	4	0.2	0.9
One	12	16	2.2	5.1
Two	0	6	0.0	1.1
Three	1	3	0.2	0.7
Four	5	2	0.9	1.3
Five	5	5	0.9	1.8
Six	0	4	0.0	0.7
Seven	2	4	0.4	1.1
Eight	20	9	3.6	5.3
Nine	6	6	1.1	2.1
Erase	2	0	0.4	0.4
Backspace	0	4	0.0	0.7
Overall:	53	63	0.8	1.8

the operator to spend a great deal of time repeating words in order to complete an entry.

A variety of ways of calculating figures of merit can be envisioned, most with a legitimate rationale. The two methods which have been chosen here are the following:

(a) For each word, the total number of errors (misrecognitions) for all operators divided by the total number of entries. For tables 1 and 2, as noted before, the total number (N) is 1,100 for each word. For table 3, N is 550.

(b) For each word, the total of errors <u>and</u> rejects divided by N.

In the first method, "percentage error" is interpretable as the rate of misrecognition, since the position is taken that rejected entries are at least <u>not</u> errors. The second method may be interpreted as the maximum of <u>unacceptable</u> responses of all kinds made by the word recognition system. The reader, of course, is free to perform whatever calculations may be desired--the raw numerical data do not change. In fact, with values of N as large as found here and numbers of "errors" as small as found here, the differences in the final percentage values vary at most by only tenths of a percentage point regardless of the formula employed.

The principal features of note in tables 1 through 3 are the very small overall error rates for all three subvocabularies and the fact that individual members of each subvocabulary were found to produce much higher than average rates of errors or rejections or both.

The message-type subvocabulary (table 1) showed the highest overall error rate, as well as the largest number of "standout" results in terms of words with unusually high error and/or reject rates. It is significant in this connection that the message-types subvocabulary was the first contact that any of the operators ever had with a word-recogniton system. It has been widely observed (reference 5, page 227, for example) that operators need to, and do, develop a knack of "talking to the box;" that is, in addition to the more general familiarization effects, such as the development of the habit of speaking at a rather uniform volume level, after several sessions of making voice entries operators tend to fall into a natural, offhand mode of pronunciation which contributes to recognition accuracy.

It must also be remembered that the data presented in table 1 represent <u>all</u> sessions and utterances, "warts and all"--including the earliest sessions where there was no retraining of troublesome words as well as the later, more nearly error-free sessions done after individual speaker/utterance problems had been detected and corrected.

In the 15 word message-types word list (table 1) there were 7 "words" which produced errors in excess of 1 percent. Two of these were composite "words," such as "accept handoff" and "report altitude." Some of the errors and a major proportion of the rejects produced by these words resulted from the difficulty of articulating them without any internal pause. This problem, however, was confined to the two speakers in one case and three in the other who, for example, produced

two-thirds of the rejects found for these expressions. The high error rate for the word "cancel" was almost entirely due to three of the eleven speakers. This, in fact, was the general case: where high error rates were found for a "word," from half to two-thirds of all the errors found occurred in the data for 1, 2, or sometimes 3 of the 11 speakers. Other speakers had no special difficulty with these words.

The second subvocabulary tested was the place-names or fixes (table 2). Seven of these words also produced error rates of over 1 percent, and again only four of them produced errors greater than 2.5 percent. As with the message types, in the extreme cases (for example "Milton," "Whitehaven," and "Benton") half or more of the errors and rejects were found in the data for only one or two of the speakers.

The last of the subvocabularies to be tested was that which consisted of the 10 decimal digits plus the 2 control words "erase" and "backspace," (table 3). The control words were included in this word list for initial testing. In entry of whole messages, as will be seen in experiment II, these control words must be made a part of every subvocabulary, since it may be necessary to correct an error or start over at any point in a message.

Here, there were three words showing an error rate greater than 1 percent, but only one of these was over 2.5 percent. The two worst cases ("one" and "eight") were again due primarily to the data from only 1 of the 11 operators. The important things about the results for this particular word list were two. First, the overall average rate of errors, just under 1 percent, confirmed results reported by the developers and other experimenters with this technology. Secondly, it was encouraging, since such a large part of message content in the languages of interest consists principally of numerical data.

ACCURACY IMPROVEMENT STUDIES.

While the recognition accuracy data for the subvocabularies of this language were impressive overall, two major considerations inspired a search for methods of improvement. In the first place, it must be remembered that the "user" here is the air traffic controller, and the principal aim of voice data entry is reduction of distraction from his or her main concern, namely continuous observation and management of the dynamic four-dimensional traffic situation. It is thus essential that detection and correction of data entry errors be brought to some irreducible minimum. The second problem is that of individual differences in voice recognition accuracy from speaker to speaker. While precision and clarity of speech are of the essence of the craft of ATC, some controllers necessarily will speak with greater uniformity than others. Thus, while the overall voice recognition error rate for the message-types subvocabulary was less than 1.5 percent, individual speaker error rates ranged from less than 0.1 percent to nearly 7 percent. With the "digits," subvocabulary, the overall average error was less than 1 percent, while the range for individuals was from zero to 2.3 percent. Similar results were obtained for the subvocabulary of fix names.

It was decided, therefore, to investigate means of error reduction and/or error correction which might be applied to the basic VIP-100

recognition algorithm. The Naval Training Equipment Center was consulted regarding some of the recognition subroutines that had been developed there for increasing recognition accuracy in their application in the ground-controlled approach trainer. These techniques as well as a variation of the same general concept which was developed for NAFEC by Threshold Technology were experimentally tried with the nonradar controller data entry language being used here. The net result, despite manipulation of the parameters of these routines, was either an increase in rejected inputs or an increase in the error rate or both. In retrospect this should not have been surprising, since the logic of these techniques was directed principally to the solution of the recognition problem where the input utterances were relatively long and largely identical with the exception of a single element.

For example, the expressions "slightly (above/below) glidepath" can be differentiated with greater accuracy if both the reference and the input images are pared down to only those parts which are nonidentical and a "second look" taken at the correspondences. This precise situation did not obtain in the word lists used here. The more common type of problem encountered was confusion of some of the pairs of words within a subvocabulary. The words "transmit" and "print strip" in the message-type list and the words "Williamsport" and "Resort" in the fix names list were among the frequent confusions. Oddly enough, even though the word "nine" (instead of "<u>niner</u>") was used in the digits word list, and nearly all errors involved the five/nine and nine/five confusions, a very high order of accuracy was obtained for both words.

In the course of trying out various alternative decision subroutines for error reduction and in reexamining our original detailed data, the experimenters were struck by some interesting features of the word durations. For every utterance in the original tests, data collection routines had recorded the word numbers and correlations for the best and second best matches and the duration (i.e., number of audio samples) of the input utterance. In the course of time normalization of utterances, the standard software had been discarding this information after use. It was an interesting curiosity of the subvocabularies that some of the errors that were common (such as Williamsport/Resort and fix/backspace/erase) were quite reliably distinguishable on the basis of utterance duration.

In the course of investigating the utility of <u>this</u> phenomenon in turn (the experimenters started collecting utterance duration data during the "training" or reference array construction mode of operation), it was further discovered that there were systematic differences in utterance duration during "training" as versus "recognition." The average duration of a word spoken repetitively during training frequently differed from the average duration of the same word spoken in a pseudorandom sequence. Since the durations differed under the two conditions, it was hypothesized that the <u>correlations</u> obtained in recognition would necessarily suffer.

The software was then modified in two ways. First, training was changed so that the speaker was presented with a pseudorandom prompting list. He or she did not simply repeat each word in the list 10 times in succession, but rather 10 times within the same

list--but seldom or never the same word twice in succession and in an unpredictable order. At the same time, the average duration of each word as well as the shortest and longest obtained during training were recorded and made a part of the reference information. The recognition decision algorithm was then changed to make use of the duration data. The basic logic is as follows:

1. The input word is digitized, time normalized, and its duration is noted.

2. The normalized feature array is compared with the reference arrays for all words in the subvocabulary, and the routine returns with the correlations for the best and second-best matches.

3. If the correlations differ by more than 40, the best match is selected as correct.

4. If the correlations differ by 40 or less, the input word duration is compared to the average (during training) duration for the first and second choice words unless the latter two durations themselves differ by less than 30 samples.

5. If the duration of the input word is closer to the reference duration of the first-choice word, it is accepted as correct.

6. If the duration of the word is closer to that of the second-choice word, the input is rejected.

7. If the two reference durations differ by 30 samples or less, the test is not made, and the first choice word is accepted as correct.

In addition to these changes in the training and recognition algorithms, a "tuneup" mode of operation was added to the basic program. In this mode of operation, the speaker puts on and adjusts the headset, adjusts the input volume setting, and then starts reading the words in the particular subvocabulary. The recognition decision word is displayed on the Tektronix terminal cathode ray tube (CRT) and just below it, the duration in samples of the word just said and the average duration of the first-choice (or recognition decision) word. If the two durations are not reasonably close (i.e., differ by more than 10 or 15 samples) for several of the words, even when repeated several times, then the headset placement and volume setting are rechecked. This "tuneup" mode is also useful for checking the effects of a cold or other speech-altering event and the need for "retraining" specific words.

Having made new training data by the pseudorandom repetition method, two of the "better" (i.e., higher overall recognition accuracy) and two of the "poorer" speakers were retested on the three subvocabularies previously used. With only one exception (fix names for one of the "better" subjects) the difference between the average duration of words in the training or reference data and the average duration of the same words under recognition conditions decreased substantially. With another similar exception, the average correlations of input words increased. That is to say, the quality of the matches between the inputs and their reference images, on the whole, improved.

As might be expected, overall errors of recognition were reduced. The

percentage error across all speakers and all three word lists went from 1.0 down to 0.35 percent. The percentage of rejects, somewhat surprisingly, went from 1.3 down to 0.8 percent. This last is surprising because it was expected that the use of duration information in the recognition decision logic would tend to increase the reject rate by rejecting some doubtful, atypical but correctly recognized (on the basis of correlation alone) spoken inputs. This was a trade we were willing to make, namely, the exchange of rejects for errors. The "cure" for a rejected entry is simple: Say it again. The cure for an error is another story entirely.

Thus, it would seem that the modified training routine alone solved most of the problems we sought to solve. In addition to this effect, the duration test in the decision logic only slightly increased the reject rate for two of the speakers on the list of fix names, while the error rate for both was reduced to zero. Indications are, overall, that use of this additional information will convert a portion of the potential errors to rejects for some talkers.

Recognition reliability or error rate improved for both the "poorer" and the "better" talkers on all three subvocabularies with only two exceptions wherein it simply remained the same. In one of these two cases, the error rate was zero under the original test conditions and, obviously, could not have been improved in any event. The improvements for the "poorer" talkers were not uniformly dramatic, but they were very impressive in most cases.

It must be admitted that in the follow-on studies reported here, we were proceeding on a "pilot-study" or "cut-and-try" basis until the very end. Thus, the final results noted just above are accounted for by a combination of variables. The training procedure was changed, the "tune-up" feature was added, and the decision logic was modified. In addition, there may have been some unknown quantity of "Hawthorne Effect" upon the "poorer" talkers who worked closely with the experimenters through the cut-and-try phase of the experimentation. The "acid test" of the objective changes should properly be made with a new sample of subjects. On the whole, however, we feel that we substantially realized our goal which was reduction of recognition error as close to the vanishing point as possible given the technology at hand.

OTHER FINDINGS.

Colds and allergies which affect the characteristics of speech were found to deteriorate recognition quality. However, for two of three speakers who among them contracted three head colds and one allergy during the test series, no serious problems were encountered. For these two speakers, it was necessary to retrain only a few of the words in the list to recover the near-perfect recognition previously found.

One speaker contracted a second cold after several weeks. It was only necessary to read into the system the training data modified for the first cold in order to achieve the same recognition quality as produced by the "normal speech" training data. Another speaker, however, despite major efforts at retraining specific words, was unable to regain a high recognition accuracy while a cold persisted.

It should be noted that the overall data for recognition of message-type entries which have already been discussed (table 1) include the error data from this speaker which accounts for approximately half the total errors encountered with this particular subvocabulary. When this speaker was not suffering from a serious cold, his results were quite comparable to those of other speakers.

Retests were also run with most of the original 12 speakers using the last (and best) set of training, or reference, data recorded during the initial reliability testing phase. Retests were made after approximately 3 months and again after approximately 6 months following the last of the original test series. Both accuracy and reject results were almost identical to those found in the initial test series.

Finally, microphone quality and placement were found to be factors of influence. While fully systematic testing of these variables was not conducted, three different (but all "noise canceling") microphone types with different mountings (one hand-held, two headset or headband) were employed at various times. The hand-held microphone was used by three of the speakers during the testing of the 15-word message-type list and accounts, in part, for the slightly lower overall accuracy rate found for that list than for the others. Careless, inconsistent, or unusual placement of microphones (e.g., at or below chin height, more than an inch from the corner of the mouth in the horizontal plane) immediately elicits a high reject rate because of loss of signal strength and can quickly be corrected by the user. The microphone used by all but one subject for the "digits" subvocabulary is directly substitutable in existing ATC operations for the carbon-type microphones required by the communications systems employed today. This microphone produced excellent results. Microphone technology has also improved since these experiments began, and some experimenters report significant performance improvements due to microphones alone.

EXPERIMENT II

RATIONALE.

It is one thing, of course, to secure a high order of recognition accuracy (greater than 99 percent for even the least proficient speakers after the incorporation of improvements in the training procedure and the recognition algorithm) for separate parts of a total language. It is quite another to generalize this performance to data entry in total real jobs-of-work. The operational tasks for which word-recognition technology was being evaluated involve the entry of whole messages, not just single words. A typical example would require the entry of an orderly sequence of utterances which convey the intention to amend a flight data store, the identity of the store or file, and the specific modification to be made. A number of examples may be found in appendices A and D. The purpose of experiment II, therefore, was to make basic comparisons between the entry of whole messages by voice as versus entry of the same messages by the keyboard method currently in operational use.

PROCEDURE.

The language chosen for test purposes was that typical of the nonradar

control position in the ATC center. Two hundred messages were constructed in two sets of one hundred each. Each hundred messages consisted of exactly:

--20 flight plan amendments. (Ten required amendment of only one data field, five amended two fields, and five amended three fields)

--16 departure messages. Eight of these included the optional altitude entry, eight did not

--14 flight plan readout requests

--13 handoff entries

--12 handoff acceptance entries

--7 flight strip printout requests

--6 weather information requests

--5 drop track (and flight plan) entries

--2 flight plan cancellations

--1 each (a) early transmission of flight plan to a terminal, (b) entry of a reported altitude from a flight without an altitude transponder, (c) track holding message, (d) track released from holding message, and (e) discrete beacon code assignment entry.

The format of the messages may be found in appendix A, and a sample of 25 of the messages may be found in appendix D.

Each set of 100 messages was written out on individual cards in narrative, descriptive form as a requirement to make an entry and not as a sequence of words to be said. The 100 cards in each instance were shuffled into more or less random order. The random sequence of messages was then transcribed onto printed sheets, 25 messages to a sheet. Appendix D contains one of the total of eight message sheets at were produced. This whole process resulted in a standardized set of messages which contained nonradar controller entries with frequencies representing those found in actual control sectors in the field. The large number of messages prevented the operators from learning messages or the sequences of operations required to enter them. Thus, every message, one by one, had to be "translated" from the descriptive form in which it was presented into a sequence of spoken words (or keystrikes) necessary to compose the message in machine acceptable form.

Each of the experimental operators was given a copy of the operators manual (appendix A) several weeks in advance of any data collection. A schedule of test sessions was arranged individually with each operator. Five operators completed the whole test series. All were ATC specialists, four male and one female. Two of the men had had extensive keyboard data entry experience in the NAFEC Enroute System Support Facility but none of this within the previous two years. One of these had also served as a test operator in experiment I and was thus more familiar with the voice entry system than the other four. Three of the operators started with and completed the voice entry system first followed by the keyboard system. Two started with the keyboard entry system and finished with the voice entry system.

Prior to data collection using the voice entry method, each operator:

1. Was given a complete demonstration of the operation of the system by the experimenter.

2. Was given preliminary training and familiarization with voice entry by creating sets of word-prints (machine "training") for the subvocabularies tested in experiment I and generating several sets of recognition data for these word lists.

3. Created an initial set of word-prints for the entire 103-word vocabulary of the voice entry language (see appendix A). This was accomplished by using the pseudorandom training method developed as previously described and,

4. Entered a set of practice messages (see appendix A).

Prior to data collection using the keyboard method, the vocabulary and syntax of the keyboard language was explained to each operator, and a set of practice messages was entered. At all times during data collection under both voice and keyboard systems, a chart was available to the operator showing the vocabulary and message structure of the language in use at the time. In addition, if the operator had difficulty formulating any message, he or she was permitted to ask the experimenter for instructions. In practice in the field, controllers are issued and/or have available for reference a pocket reminder card which describes the required format and content of the various messages in the keyboard language. It was found during the experimentation that there was a "learning" function for both languages which continued through the first two or three hundred messages entered. In fact, for the one-in-a-hundred types of messages (such as entry of reported altitude, etc., see message distribution described above) this learning function continued right through the end of the experiment. This too reflects operational experiences--the vocabulary and format of the rare types of messages are not well remembered, thus, the need for the "reminder card" noted above.

Data were collected on the entry of 100 messages (4 sets of 25) in a single session. The operator reported to the laboratory, was seated comfortably facing the Tektronix terminal display and donned the microphone headset, positioning the microphone approximately in the recommended position. The reference or word-print data for the specific operator were read into program storage and the program started in "tuneup" mode. The operator then spoke a number of words, usually digits and message-type words, while checking microphone placement and input volume setting to achieve an approximate match between input word duration (as displayed on the terminal) and reference duration (also displayed for each word).

At this time, any retraining of vocabulary words necessary was accomplished. Any words which produced recognition errors or frequent rejections during the previous session were retrained at this time, and the training data thus modified were stored for use in the ensuing data collection session. The experimenter then entered the identity of the output data file, handed the operator the first set of 25 message descriptions, and entered a "start" signal at the computer console. The operator then proceeded to translate the message descriptions into the sequence of words required to compose and enter the messages one at a time. When the operator said the last word in the last message, the experimenter entered a "stop" signal at the control console and closed the data file. The

operator was then given a short rest after which the process was repeated for the entry of the second set of 25 messages and similarly with the third and fourth. In this manner, a total of 100 messages was entered at one sitting. Generally, not more than 100 messages were entered on any given day by any one operator. In any given day, several operators would usually be scheduled. The data files were usually processed the same day as collected. In this way operator omission of whole messages, processor failures, and similar rare events could be detected and corrected at once.

An almost identical procedure was followed for keyboard data entry sessions, except that no microphone or control setting tuneup was required. The keyboard messages were entered through the standard keyboard which was an integral part of the Tektronix terminal console which also provided the operator display. In both voice and keyboard entry methods each message was displayed on the Tektronix display word by word or key by key as it was being entered by the operator. The operator could thus check the composition and accuracy of the message as it was being assembled and could make corrections or clear and reenter the message at any time prior to the Enter signal. The Enter signal was the word "go" in the voice system and the carriage return key in the keyboard system.

DATA COLLECTION AND PROCESSING.

The data collection subroutines of the computer programs developed and used here for both voice and keyboard entries performed several functions, namely:

1. The "start" signal entered by the experimenter caused a real-time clock to be read and recorded. The "stop" signal caused the clock to be read a second time. The difference, in seconds, was calculated and printed out at the control console at the end of a set of 25 messages.

2. The entry of a message type (either the first word in a voice message or the first key in a keyboard message) caused a real-time clock to be read. The entry of the "go" word or the carriage return, respectively, signaled the completion of the actual message, causing the clock to be read again, the difference to be calculated.

3. The data collection software maintained, for each message, a record of every word recognized, every word rejected, every backspace, and every erasure in the voice system and every key struck in the keyboard system. All of these together with the time elapsed between first and last entry in the message (item 2, above) were written message by message sequentially as entered into a disk store file for later processing.

Samples of complete data files, selected to illustrate all of the possible events recorded may be found in appendix E.

The data processing software was designed to perform an exhaustive tabulation of the following quantities for each data file representing the entry of 25 messages:

1. The total number of characters in completed messages. Messages with language or format errors and messages with elements omitted or

added by the operator were printed out, together with the "correct" message (from a master list) for visual analysis.

2. The total numbers of backspace and erase entries and, for the voice system only, the number of entries rejected as not recognizable.

3. The number of erroneous characters in completed messages. The master or correct message in each case was compared, character by character, with the message entered by the operator with either system. Errors in messages which were printed out because of format and language errors were visually counted and added to the totals calculated by the program. The vast majority of messages was completely processed by machine, and very little visual/ manual processing was required. Samples of processed data files may be found in appendix F.

Five of the 15 types of messages which are common at the enroute nonradar control position were "shorthand" or very brief types of entries also quite commonly entered by the radar controller. Three of these require, for example, entry of only the message type (executed by a single "quick action" button in a separate key pack at the radar position) plus the track identity (most commonly three digits entered through the numeric key pack or the alphanumeric keyboard). These are the "accept handoff," "drop track," and "flight plan readout" messages. Two others, the "handoff" and "reported altitude" messages, require a single-key message-type entry, identity, and either two digits (for handoff, to identify sector if the handoff is made to other than the expected or "normal" sector) or three digits (altitude entry). Handoff and accept handoff messages, in fact, constitute over two-thirds of all messages commonly entered by the radar controller (reference 7). Within the total of 3,000 messages entered by the five operators using each of the entry systems, there were 1,350 messages (in total) of these five types. Thus, though the total number of keystrikes required to enter each of these messages by the keyboard method in this experiment is greater by exactly one at the nonradar position, it was considered that accuracy and other measures for these five types of message considered separately would provide some indication of the relative merits of voice versus key entry at the radar position. Therefore, every set of raw data was processed twice, once to summarize performance over the whole set of 15 "high frequency" nonradar position entries and again to summarize performance for the subset of radar position entries. It should be remembered in considering the detailed results in the tables below (except for data on "translation times" where the distinction does not appear) that the results identified as "R Position" are a subset of those labeled "D Position" and are, in fact, contained within the overall summary results tabulated as "D Position" findings.

RESULTS.

Table 4 presents the numbers of errors of all kinds which were found in the messages as finally completed and entered; that is to say, the errors (either operator committed or word recognition errors or both) that remained undetected and uncorrected by the operators. There are three kinds of errors:

1. Language errors (both voice and keyboard) made by the operators.

TABLE 4. MESSAGE ENTRY ERRORS

Voice

Operator	D Position Lang.	D Position Form.	D Position Char.	R Position Lang.	R Position Form.	R Position Char.
1	6	0	14	1	0	4
2	1	0	5	1	0	1
3	2	0	7	1	0	0
4	11	0	21	3	0	3
5	7	0	47	1	0	9
Per Hundred Messages	0.9	0	3.1	0.5	0	1.3
Overall		4.0			1.8	

Keyboard

Operator	D Position Lang.	D Position Form.	D Position Char.	R Position Lang.	R Position Form.	R Position Char.
1	23	37	31	13	5	12
2	36	27	13	4	1	2
3	6	22	20	1	2	0
4	6	34	44	1	4	8
5	10	4	22	3	0	3
Per Hundred Messages	2.7	4.1	4.3	1.6	0.9	1.9
Overall		11.2			4.4	

An example here is entry of the word "drop track," or the key code "RS," when the entry should have been "cancel" or "CN," respectively, for cancellation of a flight plan.

2. Format errors made by the operator. Examples here are the omission of message field delimiters (or "punctuation" so to speak) such as spaces between fields or entry of a field designation where the format did not require one.

3. Character errors. These were single letter or digit errors in the message as entered and could arise from misrecognition of a spoken word, speaking the wrong word, or striking the wrong key.

In table 4 the errors are tabulated for all messages entered by each operator separately. Errors are also summarized as the number per hundred messages, where the total numbers of messages entered by all subjects in "D Position" messages was 3,000 and for "R Position" messages, 1,350.

An important feature to note in the voice entry results is the total absence of format errors. This is accounted for solely by the fact that the voice system, of its nature, requires a "computer in front of the computer." That is, format control is an essential part of the voice system. For example, the first entry in a message must be the message type. The syntax built into the voice entry program makes it impossible for the first entry to be anything else--whether the word said or recognized is correct or incorrect. Obviously, if the same control were applied to key entries (which, in the field, it is not--format is not inspected until a whole message is entered) format errors would be virtually impossible with key entries either.

The important comparisons to be made in table 4, considering first the nonradar or D Position messages taken as a whole, are as follows:

1. Language errors are three times as frequent with the keyboard method as with voice.

2. Single key or character errors are at least one-third more frequent with key entry as with voice.

3. If one considers the matter of errors of all possible kinds and accepts the built-in impossibility of format errors in the voice system as a real benefit, the advantage of the voice system seems quite clear. The key system produces nearly three times as many entry errors overall.

With respect to language errors, the difference can be accounted for mainly on the basis of the differences in the mental encoding or learning or remembering processes. The voice system uses a "natural" word code, while the key system uses a (necessarily) contrived or artificial code. There is further evidence of this effect in the "translation times" presented and described below. Thus, the intuitive hypothesis of an advantage to the natural language method of data entry appears to be tenable. There are fairly large individual differences from operator to operator but they appear to be much greater for the keyboard system. Operator number 3 produced nearly a third of the language errors found with the voice system, while operators 1 and 2 produced about three-fourths of the language errors in the key

system. Curiously enough, operator number 1 was one of the two who had substantial prior experience in the system support facility for enroute development which has been previously noted.

As regards single-character errors found in the voice entry results, exactly half of the character errors are found in the data for operator number 5. This operator was the only one who insisted on using a hand-held microphone for entry of the first 500 messages, finally becoming so annoyed at the results he was getting that he switched to a head-mounted microphone. Character errors for this operator thereafter virtually disappeared. If the data for this one operator are omitted from both voice and keyboard results, the character error rate per hundred messages for key entry is more than twice as high as for voice entry, 4.5 as versus 1.9 per hundred, respectively. This, of course, is a clear-cut illustration of the sensitivity of word recognition to microphone quality, technique, and positioning of which more will be said below.

It is true that not all errors are equally serious. The operational key entry system will detect many format and language errors and call them to the attention of the operator (by rejecting entry of the whole message), thus preventing their entry into the system. Single-character errors, however, are much more likely to escape detection. There is much to be said for a system which reduces the likelihood of error in the first place.

If attention is focused on the error pattern for radar position types of messages in table 4, language errors are less frequent with both entry methods (0.5 per hundred for voice and 1.6 per hundred for keyboard). If the one operator (number 1) who produced 13 language errors in 225 messages using the keyboard is overlooked as being anomalous, even this difference disappears, and the rate for key entry shrinks to 0.6 per hundred. Similarly with single key or character errors, if this one operator is not included, the character error rates become nearly identical at 1.2-1.3 per hundred with both systems. Thus, with the exception of format control, the voice system did not appear to offer any advantage with the types of messages commonly entered at the radar controller position.

It should also be pointed out here that the large number of character errors (47) attributable to operator number 5 with the voice system is precisely due to use of the hand-held microphone and disappeared when the change was made to a headset microphone. None of these results is surprising if it is remembered that this subset of five message types comprises the briefest and simplest messages in the whole possible repertoire. On the other hand, if the effects of individual operator differences are allowed to stand, the keyboard system can be seen to have produced about two to three times the number of errors of all kinds per hundred messages, even with the radar position subset.

Table 5 presents a summary of the data entry rate measurements. Here too, the results are tabulated for the whole D Position message set and separately for the R Position subset. The individual entries were calculated by dividing the total of all the entry times by the total of characters entered.

TABLE 5. DATA ENTRY RATES, CHARACTERS PER SECOND

Voice

	D Position		R Position	
Operator	Min.	Max.	Min.	Max.
1	0.9	1.6	1.0	1.3
2	1.5	1.8	1.7	2.3
3	1.2	1.6	1.2	2.0
4	1.1	1.5	1.4	2.6
5	1.1	1.9	1.1	2.1
Average	1.2	1.7	1.3	2.1

Keyboard

	D Position		R Position	
Operator	Min.	Max.	Min.	Max.
1	1.1	1.6	1.9	3.3
2	1.6	2.0	2.6	3.0
3	1.1	1.6	1.6	2.3
4	1.1	1.7	1.7	2.5
5	1.3	1.9	1.5	2.7
Average	1.2	1.8	1.9	2.8

It will be remembered that an "entry time" was measured for each message, which was the time between the first word or keystrike in the message and the last word or keystrike. The total numbers of characters were also counted, and the ratio between the two constitutes a number of characters per second.

The rates set forth in table 5 are the minimum and maximum rates for each operator under each of the conditions; i.e., the slowest and fastest sets of 100 D Position (or subsets of 45 R Position) messages out of the six sets entered by each operator. It is easily seen in table 5 that for the full set of D Position messages there was no difference of practical importance in the data rates for voice as compared to keyboard entry, the characters/second varying between 1 and 2 for both.

R Position entries, however, show a 30- to 50-percent advantage for the keyboard system. The difference in this case can be ascribed to two factors. The R Position messages, as previously noted, were "short and sweet" and consisted mostly of digits. They are simple and quick to learn and easy to remember, and the keys required for the bulk of each entry are close together. While this same simplicity applies to voice entries of these messages, the necessity of speaking each word with clear separation slows the process down and gives the key system a speed advantage. At the same time, this sort of key entry is <u>not</u> solely a touch-typing activity and necessarily distracts the visual attention of the operator. He or she <u>always</u> looks at the keys (and less often at the message being composed on the display) while entering the message. In voice operation this is not true--there is not nearly the shifting of visual attention from message list, to keyboard, to display.

These results tend to confirm the general findings of other investigators (references 8 and 9) that a <u>limited</u> capability for processing continuous (unhalting) speech would be a very important requirement in the development and application of this technology. A word recognition system which would process continuously vocalized two-, three-, or four-digit numbers would, in all likelihood, eliminate or reverse the advantage of the keyboard in the R Position message subset and produce a clear speed advantage to voice in the D Position messages.

Table 6 presents a summary of the "translation time" data. It will be remembered that two sets of time data were collected; namely, the total times to read, translate, <u>and</u> enter messages and the total times consumed in actual entry alone--either speaking or keying from first to last element of the messages. Obviously, some of the actual "translation" actually took place during the entry process.

The differences, in seconds, between the total times for each 100 messages for each operator and the sum of the "entry" times are shown in table 6. Again there were one or two outstanding individual differences; for example, operator number 1 in the keyboard entry situation and operator number 2 in the voice method. The important feature of this set of data, however, is that there is an obvious, consistent advantage to the voice system. Translation, message composition, and related elements of the data entry process were markedly facilitated by the voice entry system.

TABLE 6. MESSAGE TRANSLATION TIMES
(TOTAL SECONDS)

Voice

Session*

Operator	1	2	3	4	5	6
1	509	328	230	340	362	370
2	225	290	162	169	135	205
3	388	246	238	250	247	255
4	564	-**	297	312	274	252
5	493	358	385	370	349	226
Average	436	306	262	288	273	262

** Clock failure, data lost.

Keyboard

Session*

Operator	1	2	3	4	5	6
1	1,071	939	808	794	660	644
2	637	213	319	617	598	724
3	608	585	540	569	322	625
4	534	507	548	460	411	393
5	582	487	370	425	267	393
Average	686	546	517	573	452	556

* A session included 100 messages.

This part of the process appeared very early in the experience of the operators, after only one or two hundred messages had been entered, and the advantage remained substantial thereafter. The translation process for the keyboard system took 50 percent longer at the very beginning of the test series and quickly "rose" to the vicinity of 100 percent longer as the advantage of the voice system increased. The advantage remained even after 500 messages had been entered. The difference, at the end, was that between an average of about 2.5 seconds translation time per message with the voice system versus 5.5 seconds per message using the keyboard language.

Tables 7 and 8 summarize the numbers of "backspace" and "erase" entries made in each 100 messages by each of the operators using the two entry systems. These reflect errors which were detected and corrected prior to final entry of a completed message. The detected and corrected errors include only operator errors in the keyboard system and both operator and word-recognition errors in the voice system. Table 7 summarizes the results for the whole D Position set of messages. While the overall average numbers of corrections for the two methods appear to show about a 33-percent advantage to the keyboard system (10.8 corrections per hundred keyboard messages versus 15.7 per hundred for voice) it can also be seen that this difference is almost totally accounted for by the results from one operator (number 4). If data for this one subject are not included in calculations, the advantage is much diminished. The difference in numbers of corrections per hundred messages in favor of the keyboard is reduced by half.

A similar situation may be seen in table 8. If the results for all five operators are included in an overall rate of corrections per hundred messages, the rate is nearly twice as high for the voice system as for keyboard in the subset of R Position messages. However, operator 4, who stood out from all others in this respect under the voice entry condition for the whole class of D Position messages again accounts for nearly all of this difference. If only results for the other four operators are considered, the difference effectively disappears.

This one operator had a great deal of difficulty adapting to the requirement of the voice system that words be enunciated separately and only began to develop this skill in the last 100 messages or so. Again, much of the negative effect in this instance might be eliminated if a degree of limited continuous speech recognition emerges in this area of technology.

Finally, considering the overall results for words rejected by the voice system, there was a surprisingly large reject rate. The average 100 messages, as entered through the voice system, required 756 spoken words or utterances. (See appendix D for the count of utterances in a sample of 25 messages.)

The 3,000 messages entered by all operators taken together required 22,680 utterances. An additional 2,617 utterances were rejected in total. Thus about 10 percent of all words vocalized were rejected as unrecognizable. This is an extremely high rate compared to that of 1 percent or less found in experiment I. The causes of the high reject rate, however, are reasonably clear.

TABLE 7. ERRORS DETECTED AND CORRECTED ("BACKSPACES" PLUS "ERASURES") FOR D POSITION MESSAGES

Voice

Operator	Session* 1	2	3	4	5	6
1	17	15	29	13	14	11
2	4	7	4	15	13	12
3	8	7	4	5	5	11
4	32	51	31	41	40	15
5	10	18	18	12	5	4
Per Hundred Messages	14.2	19.6	17.2	17.2	15.4	10.6
Overall			15.7			

Keyboard

Operator	Session* 1	2	3	4	5	6
1	6	7	4	8	16	21
2	3	17	13	14	6	19
3	11	10	3	15	18	20
4	16	13	9	12	20	6
5	7	11	4	5	4	5
Per Hundred Messages	8.6	11.6	6.6	10.8	12.8	14.2
Overall			10.8			

* A session included 100 messages.

TABLE 8. ERRORS DETECTED AND CORRECTED ("BACKSPACES" PLUS "ERASURES") FOR R POSITION MESSAGES

Voice

Operator	Session* 1	2	3	4	5	6
1	8	5	9	1	7	7
2	1	0	0	2	1	1
3	3	3	0	2	1	1
4	8	16	8	13	14	5
5	9	3	1	2	5	0
Per Hundred Messages	12.8	12.0	8.9	8.9	10.7	6.2
Overall			9.9			

Keyboard

Operator	Session* 1	2	3	4	5	6
1	3	0	0	3	2	0
2	0	1	3	0	3	3
3	0	1	1	5	10	1
4	6	3	0	7	3	0
5	2	1	2	1	1	3
Per Hundred Messages	4.8	2.7	2.7	7.1	8.4	3.1
Overall			4.8			

*** A session included 45 "R Position" messages.**

Basically, all of those factors which have been found to reduce the quality of the voice input and, necessarily, the quality of the response of the word recognition system, came into play and were aggravated during the entry of whole messages in this experiment as distinguished from the entry of single words in subvocabularies. Principal among these for all of the operators was the speaking cadence problem. It is absolutely essential to provide a brief (approximately 1/10 second) "silence" between successive words. Two of the operators developed this facility very early in the testing, three others were beginning to develop it only near the end of the testing. A second major cause of this result for all subjects was the microphone quality and placement problem. Although some pains were taken to "tuneup" and check microphone placement and volume control settings at the beginning of each voice entry session, the head-mounted microphones would move or be moved by operators and voice amplitude would change over time. Attempts were made during testing to reduce the reject problem by "retraining" or producing new word-prints for words which were giving high reject rates, but by and large this approach was nonproductive. The one subject previously noted who used a hand-held microphone for five out of the six sessions with the voice system produced a significant fraction of the total count of rejects. When this operator changed to a headset microphone, the reject rate dropped by a factor of five.

Here once more, development of even the limited continuous-speech capability previously mentioned would reduce the reject problem markedly and quite probably to acceptable proportions. The operators who frequently forgot to space out the spoken digits separately would be rewarded, most the time, by three or four correctly recognized digits rather than held back and forced to repeat by one or more reject signals.

OTHER FINDINGS.

During the experimentation described above, NAFEC engineers working on the flight service station (FSS) improvement program developed a practical and inexpensive device which provided the capability of encoding human speech in digital form, storing the digital representation, and reproducing it at will. A version of this device was constructed and attached to the voice recognition system. The purpose of this effort was to permit trial and, if practical, the implementation of audio verification of messages entered--either by voice or by keyboard. Software was also developed to store, filter, retrieve, and reproduce the digitized speech, and a number of voice files were created representing the data entry language of the nonradar controller. In operation, this system functioned as follows:

1. The data entry operator spoke the sequence of words necessary to compose a complete message.

2. Immediately after the operator spoke the word "go" which signaled completion of the message, the sequence of words which the system had recognized was repeated back over a loudspeaker to the operator. These were not the operators own words but rather the sequence of words that the recognition device "thought" he had said, now retrieved from a disk store made from the speech of another person, reconstituted into audio form and output through a speaker.

It was found in trials by a number of operators that, given the data entry language being used, the process of audio verification was slow and too transitory to be conveniently used. The output rate is controllable, but it was found that an output rate faster than the original input rate (successive words separated by less than approximately 1/10 second between words) was too fast for complete comprehension by the operator. Even this fairly rapid rate, however, made the message verification process much slower than visual checking of the message as it was composed on the operator's display. In addition, it was found difficult to "visualize" the sequential position of an error detected in the once spoken and now "gone" audio output. Thus, correction of errors detected was more difficult than it was in the case of visual verification. As a result, further application of auditory feedback was suspended prior to experiment II. This capability, however, is expected to have utility in other applications of word recognition technology such as the development of automated pseudo-pilots for ATC simulations and controller training.

ANALYSIS

The results of the experiments reported here present a mixed picture regarding the potential operational value of word recognition technology, at its present state of development, as a substitute for the keyboards presently used as a means of traffic control data entry.

In general, voice data entry is at least as effective overall as current keyboard methods. In several respects it offers demonstrable advantages. Because of its use of natural language, voice entry produced fewer language errors by a factor of three and saved 5 minutes of translation time per 100 messages entered as compared to keyboard entry. Because of the high rate of word recognition accuracy, voice entry resulted in 33 to 50 percent fewer single character errors (equivalent to mis-struck keys) in completed, entered messages than keyboard entry. Because of the sheer presence of the word recognition equipment which included a mini-computer with software format control, the voice system absolutely prevented errors in message format which were rather common (up to 4 per hundred messages) with the keyboard system. The substantial saving in translation time previously noted also reflects an important reduction in distraction of thought and attention to the data entry task. While the gains to safety and efficiency in the traffic control process resulting from all of these advantages would be very difficult to quantify, they are clearly nontrivial.

On the other hand, the keyboard entry system as currently employed in the operational enroute control system showed two advantages over the voice system. In the subset of radar controller messages entered by both systems, there was a 50 percent higher data entry rate with the keyboard than the voice system. While this advantage to the keyboard did not appear in the nonradar controller messages (and was, in fact, neutralized by a faster entry rate for the additional 10 types of messages in the nonradar set), it is an advantage of some import to the evaluations reported here. The principal, if not sole, reason for this keyboard advantage is the single-word-at-a-time limitation of the recognition technology employed in these experiments. This fact, in

turn, points strongly to a need for development of at least a limited continuous-speech recognition capability.

The other advantage of the keyboard entry method employed in this study was that less effort was exerted by the operators in the detection and correction of errors during the message composition and keying process. A net result, of course, was the persistence of more undetected errors in the completed keyboard messages which has already been cited as an advantage of the voice system. However, the fact that the voice system elicited more corrections (i.e., backspaces and erasures) bears on the responsiveness or convenience of use of the voice system as currently configured. Enunciating separate words close together in time will almost always result in rejection of the words or an error of recognition or both. The rejection of an input is signaled by an audio tone or "beep," calling the attention of the operator to the display. As has been seen in the results of these experiments, operators will most often correct errors resulting from improper input cadence. Rejections due to this same cause are recognized as such with the result that the operator then pays special attention to adjusting input cadence to overcome the problem. There is no doubt, however, that the necessity to adapt to this limitation of word recognition systems is a source of annoyance to operators, even in the relatively uncomplicated environment of the laboratory. As with a number of other aspects of the performance of the operator/machine voice entry system (such as error rates, data entry rates, message translation times) all of which tend to be interrelated, this factor could not be expected to improve in the operational environment. It is much more likely that overall performance of both the voice system and the keyboard system used in these experiments found in the laboratory would deteriorate in the operational setting where they would be enmeshed in, and interact with, the other task components of the total air traffic controller job.

Even as these experiments were being conducted, the technology of voice input to computers has been advancing. There have been some improvements in microphone design. There are other developments just emerging at this writing (reference 10) which appear to bear some potential for improving the operator interface--perhaps requiring less adaptation of the speech habits of the speaker, less sensitivity to microphone positioning and technique, and simplified methods of "tuning" the voice system to the individual operator. If some of these improvements can be realized and verified, they may greatly reduce or eliminate limitations of voice entry in the applications investigated here.

Meanwhile, the results of these experiments indicate that word recognition technology should be given serious consideration as an alternative to keyboard data entry in ATC applications of the reasonably near future. Specific applications which appear capable of benefiting from the advantages of voice entry are the radar control positions in enroute control where the reduced distraction of attention could be a valuable safety feature and control tower cab operations where the data entry elements of the jobs are less onerous than in enroute control and where keyboards and other manual controls

cannot be made continuously and conveniently accessible to traffic controllers.

Finally, it seems evident that the aviation operational and evaluation community should continue to maintain contact with and contribute to this area of rapidly developing technology.

CONCLUSIONS

1. Word recognition technology at its current state of development has demonstrable advantages in accuracy, simplicity, and convenience over existing keyboard methods of data entry in ATC applications.

2. Present-day word recognition has some definite limitations, such as sensitivity to operator characteristics and habits and a demonstrated lack of advantage in data entry rates in the traffic control applications tested.

RECOMMENDATIONS

1. Word recognition technology should be given serious consideration as an alternative to keyboard data entry in:

a. Applications where the data entry job element is a source of distraction of visual and mental attention when performed with keyboards.

b. Applications where operator access to keyboards is limited.

2. The development of improvements in speech processing technology should be awaited and evaluated before adoption in any major upgrading of existing air traffic systems.

REFERENCES

1. Eldredge, D., Before and After ARTS III Level of Automation at Houston Terminal, Project 012-603-010, U.S. DOT, FAA, NAFEC, Data Report, June 1973.

2. Eldredge, D., et al; Before and After ARTS III Automation at Boston Terminal, Project 012-603-010, U.S. DOT, FAA, NAFEC, Data Report, June 1973.

3. Moser, H. M. and Dunham, R. E., An Examination of the Spoken Vocabulary Used in Air Traffic Control, Air Force Cambridge Research Center, TN 59-73, Bedford, Mass., November 1959.

4. Turn, R., Hoffman, A., and Lippiat, T., Military Applications of Speech Understanding Systems, ARPA R-1434, Washington, D.C., June 1975.

5. Breaux, R., Curran, M., and Huff, E. M. eds., Proceedings of Symposium: Voice Technology for Interactive Real-Time Command/Control Systems Application, Naval Training Equipment Center, Naval Air Development Center, National Aeronautics and Space Administration, Moffett Field, California, December 1977.

6. Yasaki, E. K., "Voice Recognition Comes of Age," Datamation, August 1976.

7. Kennedy, R. A. Analysis of Los Angeles Center Sector Input and Output Messages, MWP-11018, MITRE Corp., May 1975.

8. Grady, M. W., and Hicklin, M., Use of Computer Speech Understanding in Training: A Demonstration Training System for the Ground Controlled Approach Trainer, Final report, NAVTRAEQUIPCEN 74-C-0048-1, U.S. Navy, December 1976.

9. Tsuruta, S. "DP-100 Voice Recognition System Achieves High Efficiency," Japan Electrical Engineering, Tokyo, July 1978.

10. Porter, J. E., Grady, M. W., Hicklin, M. B., and Lowe, L. F., Use of Computer Speech Recognition in Training: A Preliminary Investigation of a Limited Continuous Speech Recognition Capability. NAVTRAEQUIPCEN 74-C-48-2, U.S. Navy, June 1977.

APPENDIX A

OPERATOR'S MANUAL

The voice entry language for the NAS enroute flight data control position closely parallels that presently employed for keyboard entry at the same position. Thus, every message composed has a rigid format. The specific words chosen by the user to represent keystrikes or combinations of keystrikes are not sacred; however, as Humpty Dumpty said, a word can be chosen by the user to mean "precisely what you intend it to mean, nothing more; nothing less," since the user "is the Master here." The specific words suggested by the "training" sequence of the voice entry system have been experimentally tested, though, and have been found comparatively easy to use and remember as well as reliably recognized by the system. "Niner," for example, is a better choice than "nine."

The voice system functions in two modes: Training and Operation. At any given time, the machine expects to hear from only one particular person whom it has heard before. That is, it is necessary for anyone who is to use the system to have trained <u>it</u> beforehand. Training the system consists simply in the user saying each word in the total vocabulary several times. By this means the system establishes a composite reference image of the way the particular speaker vocalizes each word in the vocabulary. Different speakers, of course, have different voice pitch characteristics, different vocal tracts and different personal pronunciations and dialects. In general, this training process need only be done <u>completely</u> one time. Experience with use of the system will reveal occasional confusions (a common one is 8 for 3) which can be reduced or eliminated by retraining; i.e., creating new reference images, just for those words.

It is generally most undesirable for the speaker to adopt an unnatural pronunciation when training or retraining a word. While this works for the moment (for example stressing the "t" sound in "eight") to eliminate erroneous recognitions by the system, it immediately fails (usually resulting in rejects or complete nonrecognition of utterances of the particular word) as soon as the speaker reverts to habitual, natural pronunciation. The exact form of utterance used by any single speaker, as noted above, is not terribly important as long as it is always used. For example, you may choose to use the expression "amendment" instead of the suggested "amend" to produce the ASCII translation AM, but you cannot say "amend" sometimes and "amendment" other times and still secure accurate recognition.

The reference recognition patterns for each speaker are stored magnetically--on tape or disk, for example--and in order to ready the system for any specific person at any time the only action required is to read into core storage the reference data for that speaker.

In the operational mode, the system "sits waiting" for rigidly formatted messages, even as the present keyboard entry system. Messages cannot be composed free form or in ideomatic continuous speech. The first

utterance expected by the system is one of the words specifying a kind of message, such as AMEND or ACCEPT HANDOFF (see appended lists of "parts of speech"). Recognition of one of these "words" produces a two-character code on the display unit, such as AM or HO, followed by a space. In most instances, the next part of the message must be a three-decimal digit number (the computer I.D.) to identify the flight or flight data store to which the message applies. The three digits must be spoken separately, word by word, for example "three, six, eight" (not threesixtyeight"). The third digit is automatically followed by a space. The next following entry expected by the system depends on the kind of message selected by the first utterance--AMEND, for example, followed by the three digit I.D., thereafter expects the name of a flight plan data field such as SPEED or ALTITUDE (see list of field names) followed by an entry appropriate to that data field (e.g., "four, two, five" or "three, seven, zero"). Spacing, when required, is automatically supplied. After the modified data have been entered, the system expects either the name of another data field or one of the control words: GO (enter), BACKSPACE, or ERASE. The "GO" word (you could say "enter" or "finished" or almost any other word as long as you trained it that way and always said it that way) causes the message to be transmitted from the data entry system to the processor as a completed message. The functions of BACKSPACE and ERASE are practically self-explanatory but are also described in the attached notes and reference tables.

Besides rigid adherence to the message format rules (see tables attached) the user requires only four cautions:

1. Upon taking over the operating position, it is necessary to position the microphone with some care at approximately the same distance and direction from the mouth at all times.

2. Set the input volume control to the same setting found during training to be the best for the speech loudness habits of the particular talker.

3. Say the words naturally, without pauses within a "word" even if they are actually phrases, such as: DISCRETECODE.

4. Pause briefly (about 1/10 second) between "words," as: DISCRETECODE, ONE, THREE, ONE, etc. to allow the machine to separate and translate each word in the message.

Certain types of messages currently possible with the existing key entry system are not (at present) possible at all with the voice system. Flight data field 10, route of flight, for example is not programed. There are two connected reasons for this. One is that an exhaustive list of fixes, intersections and airways would be extremely large and the time-to-train, time-to-search, and recognition accuracy would suffer very seriously. The second reason, of course, is that entry of total flight plans or extended routes-of-flight at controller operating positions is mercifully rare and generally undesirable in any event. Another case of (at least for the present) difficult voice entry is alphanumeric identity (airline flight, or military name/# or airframe tail number). As can be imagined, this would be necessarily complex, especially since airline names, military code-names, plus the obvious phonetic alphanumerics would all be involved. The method we have

provided for entry of aircraft types (explained in detail in the attached tables and notes) is rather similar to that which is required for alphanumeric I.D. and is indeed rather cumbersome.

The fact of the matter at this point is that the whole phonetic alphabet is part of the vocabulary of this system. (It should be noted here that this can be whichever of the many phonetic alphabets you are comfortable with, not necessarily the items displayed in the prompting words for initial system training.) Furthermore, the logic for assembling strings of letters and numbers is inherent in the program used for message assembly. Thus, with no great difficulty, the system could be modified to include the capability of a full "voice keyboard" with which any message whatever could be (laboriously) composed "voice key" by "voice key." At this juncture, this is unnecessary for purposes of basic experimentation with the system. In the longer run, it would seem preferable to modify the operational procedure than to require extensive keypunching in any event.

VOICE DATA ENTRY: D-CONTROLLER VOCABULARY

DIGITS PART OF SPEECH

WORD NO.*	SPOKEN WORD	PRINT WORD DISPLAYED
0	ZERO	0
1	ONE	1
2	TWO	2
3	THREE	3
4	FOUR	4
5	FIVE	5
6	SIX	6
7	SEVEN	7
8	EIGHT	8
9	NINER	9
	CONTROL WORDS (SEE ALSO #102 ERASE)	
10	BACKSPACE	
11	GO	(ENTER)

*For retraining and other purposes.

MESSAGE KINDS

WORD NO.*	SPOKEN WORD	PRINT WORD DISPLAYED
12	AMEND	AM
13	CANCEL	CN
14	CORRECTION	CR
15	DEPARTURE	DM
16	DISCRETECODE	DQ
17	READOUT	FR
18	ACCEPTHANDOFF	HO
19	HANDOFF	HO
20	DROPTRACK	RS
21	PRINTSTRIP	SR
22	HOLD	HM
23	RELEASE	HM
24	REPORTALTITUDE	RA
25	WEATHER	WR
26	TRANSMIT	RF

*For retraining and other purposes.

FLIGHT DATA FIELD NAMES

WORD NO.*	SPOKEN WORD	PRINT WORD DISPLAYED
27	TYPE	03
28	QUALIFIER	03
29	BEACONCODE	04
30	SPEED	05
31	FIX	06
32	TIME	07
33	ALTITUDE	08
34	IDENT	02

*For retraining and other purposes.

FIXES

WORD NO.*	SPOKEN WORD	PRINT WORD DISPLAYED
35	WILLIAMSPORT	IPT
36	SELINGSGROVE	SEG
37	MILTON	MIP
38	HAZELTON	HZL
39	WILKESBARRE	AVP
40	EASTTEXAS	ETX
41	LAKEHENRY	LHY
42	TOBYHANNA	TSD
43	ALLENTOWN	ABE
44	STILLWATER	STW
45	BENTON	7QB
46	SWEETVALLEY	7EV
47	LOPEZ	7LE
48	SNYDERS	7YX
49	SLATINGTON	7ZO
50	WHITEHAVEN	9WT
51	RESORT	9ZT
52	PENNWELL	7PW
53	HUGUENOT	HUO
54	SOLBERG	SBJ
55	FREELAND	7FE

*For retraining and other purposes.

AIRCRAFT TYPE NAMES

WORD NO.*	SPOKEN WORD	PRINT WORD DISPLAYED
56	BOEING	B
57	DOUGLAS	DC
58	LOCKHEED	L
59	CONVAIR	C
60	VICKERS	VC
61	NORD	N
62	BRITISH	BA
63	GENERAL	--
64	MILITARY	--
65	DEHAVILLAND	DH

*For retraining and other purposes.

PHONETIC ALPHA

WORD NO.*	SPOKEN WORD	PRINT WORD DISPLAYED
66	ALPHA	A
67	BRAVO	B
68	CHARLIE	C
69	DELTA	D
70	ECHO	E
71	FOXTROT	F
72	GOLF	G
73	HOTEL	H
74	INDIA	I
75	JULIET	J
76	KILO	K
77	LIMA	L
78	MIKE	M
79	NOVEMBER	N
80	OSCAR	O
81	PAPA	P
82	QUEBEC	Q
83	ROMEO	R

*For retraining and other purposes.

PHONETIC ALPHA (Continued)

WORD NO.*	SPOKEN WORD	PRINT WORD DISPLAYED
84	SIERRA	S
85	TANGO	T
86	UNIFORM	U
87	VICTOR	V
88	WHISKEY	W
89	XRAY	X
90	YANKEE	Y
91	ZULU	Z

*For retraining and other purposes.

"QUALIFIERS"**

WORD NO.*	SPOKEN WORD	PRINT WORD DISPLAYED
92	DISCRETE	/U
93	DISCRETE DME	/A
94	DME	/D
95	NONDISCRETE	/T
96	NONDISCRETE DME	/B
97	TRANSPONDER	/X
98	TRANSPONDER DME	/L
99	TACAN	/M
100	TACAN 64	/N
101	TACAN DISCRETE	/P

**These expressions are to be said as all one word such as "discrete dee em ee," even though printed here and on the training display as separate words.

CONTROL WORD
(SEE ALSO #10 BACKSPACE and "11 GO)

102	ERASE	Erases Entry

*For retraining and other purposes.

CONTROL WORDS

GO----------------------------------Momentarily prints (ENTER) on display, then clears the screen. The message, including any backspacing, is recorded for data collection purposes, together with the time, in seconds, between the selection of a message kind entry and the GO entry.

ERASE-------------------------------Clears whole message, awaits a "Message Kind" entry.

BACKSPACE---------------------------Removes last spoken entry, awaits replacement from the same subset of words.

Note: Backspace, due to the storage characteristics of the Tektronix display, erases the whole screen then rewrites the message all but the last entry made by voice. If this (i.e., the entry backspaced out) was a single character in a string of digits (as in a time, altitude, speed, beacon code) or alpha/digits (as in General or Military types), only the one character will disappear, and the machine will await another number or letter. If the last entry (i.e., "word") was a data field (e.g., speed = 05) or a fix (e.g., Williamsport = IPT) etc., the whole string, such as 05 or IPT will be removed, and the machine will await another entry from the same class of words as the word deleted (e.g., "altitude" instead of "speed" or "Allentown" instead of "Williamsport").

VOICE DATA ENTRY: D-CONTROLLER LANGUAGE STRUCTURE

KIND OF MESSAGE	SEQUENCE			
AMEND	3 DIG IDENT,	DATA FIELD NAME	DATA ENTRY FOR FIELD	GO*
CORRECTION		"	"	"

DATA FIELD NAMES	VOICE ENTRIES REQUIRED
Type	See Below
Beacon Code	4 Octal Digits
Speed	3 Decimal Digits
Fix	Place Name
Time	4 Decimal Digits
Altitude	3 Decimal Digits
Qualifier	See List of Qualifiers and Note 2
Ident	6 Alphanumerics

Note 1: After a "field name" and appropriate entries for that field have been entered, the system will accept <u>another</u> field name (plus proper entries) and yet another, etc., without limit, OR it will accept an ERASE command, a BACKSPACE command, or a GO (ENTER) command. For detailed description of ERASE and BACKSPACE, see attached Note on the subject.

<u>FOR "TYPE" ENTRIES, ALWAYS SAY:</u>

MFG NAME,	2 or 3 A/N,	Name a Qualifier	
or MILITARY,	4 A/N,	" "	"
or GENERAL,	4 A/N,	" "	

IF YOU SAY:	YOU'LL SEE:	THEN SAY:
Boeing	B	3 A/N e.g. 707
British	BA	2 A/N e.g. 11
Vickers	VC	2 A/N e.g. 10
Lockheed	L	3 A/N e.g. 011
Nord	N	3 A/N e.g. 026
deHavilland	DH	2 A/N e.g. C6
Douglas	DC	2 A/N e.g. 10
Military	--	4 A/N e.g. C131
General	--	4 A/N e.g. PA13

***The last entry in a message, of course, is always GO if all is correct. However, even at this point (as at any other) it is possible to ERASE or BACKSPACE.**

TO ENTER A "TYPE," YOU MUST ALWAYS ADD ONE OF THE EQUIPMENT QUALIFIERS:

IF YOU SAY:	YOU'LL SEE
Discrete	/U
Discrete DME	/A
DME	/D
Nondiscrete	/T
Nondiscrete DME	/B
Transponder	/X
Transponder DME	/L
TACAN	/M
TACAN64	/N
TACANDiscrete	/P

Note 2: If you wish to enter an amendment to the QUALIFIER part of the "type" field alone, you need only name the data field "QUALIFIER" then name one of the qualifiers above.

FINALLY, YOU MAY SAY "GO" (to ENTER), BACKSPACE (if you wish to change or correct an error of entry or of recognition) or "ERASE,"

OR, YOU MAY NAME ANOTHER DATA FIELD AND CONTINUE AS BEFORE.

EXAMPLES:

Say: Amend, three, three, one, altitude, three, two, zero, go
See: AM 331 08 320 (screen erases at GO)

Say: Amend, two, zero, five, type, Boeing, seven, zero, seven, discrete-deemee, go
See: AM 205 03 B707 /A (screen erases at GO)

Say: Correction, speed, two, two, zero, type, military, Charlie, one, three, one, TACAN, go
See: CR 05 220 03 C131 /M (screen erases at GO)

Say: Correction, qualifier, nondiscrete, go
See: CR 03 /T (screen erases at GO)

KIND OF MESSAGE	SEQUENCE			
REPORTALTITUDE	3 DIG IDENT,	3 DECIMAL DIG (ALT),		GO*
DISCRETECODE	3 DIG IDENT,	4 OCTAL DIG (CODE),		GO*
HANDOFF	3 DIG IDENT	2 DIG (SECTOR)		GO*
DEPARTURE	3 DIG IDENT,	4 DEC DIG (TIME)	(3 DIG ALT.)	GO*

These messages consist of the message kind followed by all digits. The altitude for departures is optional and must be preceded by the word "altitude" where made, e.g., "DEPARTURE, THREE TWO ZERO, ONE FOUR TWO FIVE, ALTITUDE, TWO ONE ZERO."

KIND OF MESSAGE	SEQUENCE	
DROPTRACK	3 DIG IDENT,	GO*
PRINTSTRIP	"	"
ACCEPTHANDOFF	"	"
READOUT	"	"
CANCEL	"	"

These messages are all identical except for the first word, the kind of message.

KIND OF MESSAGE	SEQUENCE			
HOLD	"	4 DIG (TIME)	NAME (FIX)	GO*
RELEASE	"	4 DIG (TIME)		GO*
TRANSMIT	"	NAME (FIX)		GO*

These messages require entry of a four digit time, or a one word place name (FIX) or both in addition to the message kind and the identity of the flight.

KIND OF MESSAGE	SEQUENCE	
WEATHER	NAME (FIX)	GO*

This and CORRECTION (above) are the only kinds of messages that are not immediately followed by identity.

*This last entry in a message, of course, is always GO if all is correct. However, even at this point (as at any other) it is possible to ERASE or BACKSPACE.

EXAMPLES OF VOICE ENTRY

Note: When you say "GO" at the end of a message, (ENTER) will be written briefly, then the whole message will disappear.

Modify assigned altitude of track #221 to level 370:

Say: AMEND; two, two, one; altitude; three, seven, zero; GO
SEE: AM 221 08 370 (ENTER)

Correct rejected message in speed data field to 420:

Say: CORRECTION; speed; four, two, zero; GO
See: CR 05 420 (ENTER)

Modify aircraft type and qualifier, track #397, to Boeing 707, discrete code with DME:

Say: AMEND; three, niner, seven; type; Boeing, seven, zero, seven; discretedeemee; GO
See: AM 397 " 03 B707 /A (ENTER)

Correct rejected message, qualifier only, to discrete code transponder:

Say: CORRECTION; qualifier; discrete; GO
See: CR 03 /U (ENTER)

Handoff, to sector 12, track #424:

Say: HANDOFF; four, two, four; one, two, GO
See: HO 424 12 (ENTER)

Accept the Handoff of track #337:

Say: ACCEPTHANDOFF; three, seven, seven; GO
See: GO 377 (ENTER)

Note: PRINTSTRIP (SR), READOUT (FR), CANCEL (CN), and DROPTRACK (RS) have the same format except for the code for the kind of message.

Enter departure message for track 131, time 2025, altitude 175

Say: DEPARTURE; one, three, one; two, zero, two, five; altitude; one,seven,five; GO
See: DM 131 2025 08 175 (ENTER)

Enter reported altitude of 350 for track #952

Say: REPORTALTITUDE; niner, five, two; three, five, zero; GO
See: RA 952 350 (ENTER)

Enter discrete code of 2200 for track #756:

Say: DISCRETECODE; seven, five, six; two, two, zero, zero; GO
See: DQ 756 2200 (ENTER)

Enter hold message for track 333, time 1445 at Williamsport:

Say: HOLD; three, three, three; one, four, four, five; Williamsport; GO
See: HM 333 1445 IPT (ENTER)

Enter a release (hold) message at 1500 hours for track #333:

Say: RELEASE; three, three, three; one, five, zero, zero; GO
See: HM 333 1500 (ENTER)

To force a flight plan (e.g. #123) to an ARTS terminal (e.g. Allentown) prior to the scheduled time (e.g., early flight):

Say: TRANSMIT; one, two, three; Allentown; GO
See: RF 123 ABE (ENTER)

To obtain a weather readout, for Williamsport

Say: WEATHER; Williamsport; GO
See: WR IPT (ENTER)

Amend identity of track #416 to American 142:

Say: AMEND; four, one six; IDENT; Alpha, Alpha zero, one, four, two; GO

See: AM 416 02 AA0142 (ENTER)

EXAMPLES OF AIRCRAFT TYPES

COMMERCIAL

B707, B727 B747, B737

DC09, DC10, DC81

L101, L49C, L49E, L188

DH06, DH64

VC07, VC09

CV88, CV58, CV99

BA11, BA10, BA15

MILITARY

C135, C131, C05A

F102, F11D, F11F

B058, B052, B057

KC97, DC35

GENERAL

AC68

BE33, BE55, BE80

C180, C185, C310, C340

DH03

M020, M021

N265, N40A, NA16

PA22, PA28, PAZT

T039

Note: It is realized that the four-character system in use for voice entry will not easily accommodate all type designators. It is employed here for test use only.

APPENDIX B

SAMPLE PSEUDORANDOM WORD LIST

ZERO	THREE	TWO	SIX
ONE	SEVEN	ERASE	THREE
TWO	BACKSPACE	SIX	BACKSPACE
THREE	FOUR	ONE	EIGHT
FOUR	EIGHT	NINE	FIVE
FIVE	ZERO	FIVE	TWO
SIX	FIVE	ZERO	ONE
SEVEN	NINE	EIGHT	ERASE
EIGHT	ONE	FOUR	SEVEN
NINE	SIX	BACKSPACE	FOUR
ERASE	ERASE	SEVEN	ZERO
BACKSPACE	TWO	THREE	NINE
ONE	FOUR	ERASE	
THREE	NINE	NINE	
FIVE	TWO	SEVEN	
SEVEN	FIVE	FIVE	
NINE	ERASE	THREE	
ERASE	THREE	ONE	
ZERO	SIX	BACKSPACE	
TWO	BACKSPACE	EIGHT	
FOUR	SEVEN	SIX	
SIX	ZERO	FOUR	
EIGHT	EIGHT	TWO	
BACKSPACE	ONE	ZERO	
TWO	BACKSPACE	THREE	
FIVE	ERASE	ERASE	
EIGHT	NINE	FIVE	
BACKSPACE	EIGHT	TWO	
THREE	SEVEN	NINE	
SIX	SIX	FOUR	
NINE	FIVE	ONE	
ZERO	FOUR	EIGHT	
FOUR	THREE	ZERO	
SEVEN	TWO	SEVEN	
ERASE	ONE	BACKSPACE	
ONE	ZERO	SIX	

APPENDIX C

SAMPLE OF RAW DATA, EXPERIMENT I

SAMPLE OF PROCESSED DATA, EXPERIMENT I

SAMPLE RAW DATA
EXPERIMENT I

FIRST CHOICE WORD NUMBER
FIRST CHOICE "CORRELATION"
SECOND CHOICE WORD NUMBER
SECOND CHOICE "CORRELATION"
DURATION OF WORD SPOKEN

CORRELATION LESS THAN 80, THEREFORE REJECTED

REJECTED, THOUGH CORRECT

+00000	+00231	+00007	+00136	+00176
+00001	+00232	+00005	+00107	+00094
+00002	+00240	+00000	+00108	+00111
-00003	+00070	+00010	+00064	+00214
+00004	+00216	+00001	+00092	+00133
+00005	+00212	+00001	+00137	+00147
+00006	+00179	+00010	+00109	+00222
+00007	+00205	+00006	+00113	+00165
+00008	+00141	+00002	+00109	+00083
+00009	+00156	+00011	+00135	+00145
+00010	+00241	+00003	+00132	+00215
+00011	+00145	+00010	+00095	+00277
+00001	+00193	+00005	+00095	+00101
+00003	+00178	+00002	+00125	+00156
+00005	+00181	+00009	+00135	+00134
+00007	+00168	+00006	+00104	+00160
+00009	+00142	+00011	+00091	+00154
+00010	+00274	+00006	+00156	+00221
+00000	+00228	+00007	+00126	+00175
+00002	+00249	+00003	+00142	+00124
+00004	+00246	+00001	+00132	+00144
+00006	+00157	+00010	+00091	+00209
+00008	+00235	+00003	+00126	+00091
+00011	+00165	+00010	+00100	+00266
+00002	+00242	+00000	+00121	+00131
+00005	+00210	+00009	+00130	+00164
+00008	+00189	+00003	+00113	+00085
+00011	+00184	+00010	+00116	+00280
+00003	+00160	+00002	+00109	+00138
+00006	+00147	+00010	+00122	+00225
+00009	+00156	+00011	+00079	+00140
+00000	+00222	+00002	+00106	+00180
+00004	+00225	+00009	+00100	+00161
+00007	+00210	+00000	+00113	+00157
+00010	+00287	+00003	+00170	+00210
+00001	+00192	+00004	+00121	+00120
+00003	+00246	+00010	+00159	+00170
+00007	+00234	+00006	+00106	+00175
+00011	+00176	+00010	+00105	+00266
+00004	+00230	+00009	+00130	+00153
+00008	+00152	+00003	+00095	+00087
+00000	+00212	+00007	+00098	+00183
+00005	+00199	+00009	+00169	+00162
+00009	+00166	+00011	+00097	+00144
+00001	+00202	+00005	+00127	+00109
+00006	+00190	+00010	+00088	+00188
+00010	+00262	+00003	+00165	+00202
+00002	+00210	+00000	+00116	+00118
+00004	+00215	+00009	+00101	+00146
+00009	+00182	+00011	+00089	+00151
+00002	+00215	+00000	+00115	+00117
+00005	+00198	+00009	+00174	+00149
+00010	+00281	+00003	+00160	+00213
+00003	+00260	+00010	+00161	+00130
+00006	+00211	+00010	+00117	+00195
+00011	+00166	+00003	+00046	+00200
+00007	+00243	+00006	+00114	+00166
+00000	+00206	+00002	+00082	+00188
+00008	+00175	+00003	+00105	+00079

79-20-C-1

SAMPLE RAW DATA
EXPERIMENT I (Continued)

+00001	+00197	+00004	+00136	+00110	
+00011	+00132	+00008	+00054	+00196	
+00010	+00261	+00006	+00155	+00221	
+00009	+00199	+00011	+00114	+00147	
+00008	+00210	+00011	+00106	+00099	
+00007	+00233	+00009	+00131	+00161	
+00006	+00209	+00010	+00155	+00198	
+00005	+00185	+00009	+00155	+00137	
+00004	+00231	+00009	+00120	+00152	
+00003	+00243	+00010	+00168	+00122	
+00002	+00238	+00007	+00120	+00128	
+00001	+00225	+00005	+00144	+00117	
+00000	+00186	+00007	+00064	+00183	
+00002	+00264	+00000	+00136	+00132	
+00010	+00283	+00003	+00174	+00204	
+00006	+00216	+00010	+00130	+00189	
+00001	+00218	+00005	+00125	+00125	
+00009	+00219	+00005	+00129	+00169	
+00005	+00170	+00009	+00146	+00136	
+00000	+00213	+00002	+00113	+00177	
+00008	+00167	+00003	+00105	+00082	
+00004	+00217	+00001	+00114	+00149	
+00011	+00185	+00008	+00086	+00239	
+00007	+00215	+00006	+00110	+00160	
+00003	+00249	+00010	+00158	+00109	
+00010	+00235	+00003	+00130	+00203	
+00009	+00144	+00005	+00138	+00081	
+00007	+00225	+00006	+00105	+00159	
+00005	+00193	+00009	+00159	+00140	
+00003	+00271	+00010	+00185	+00135	
+00001	+00226	+00005	+00117	+00106	
+00011	+00191	+00008	+00089	+00248	**REJECTED DURATIONS MISMATCHED**
+00008	+00183	+00003	+00121	+00082	
-00011	+00120	+00007	+00107	+00119	**ERROR (BEST CHOICE WAS "ERASE"). WAS ACTUALLY REJECTED.**
+00004	+00226	+00009	+00132	+00145	
+00002	+00208	+00008	+00118	+00105	
+00000	+00211	+00002	+00092	+00183	
+00003	+00266	+00010	+00165	+00124	
+00010	+00271	+00003	+00146	+00209	
+00005	+00222	+00009	+00190	+00144	
+00002	+00212	+00008	+00108	+00106	
+00009	+00236	+00004	+00105	+00154	
+00004	+00233	+00009	+00088	+00129	
+00001	+00223	+00007	+00130	+00126	
+00008	+00227	+00003	+00151	+00079	
+00000	+00221	+00002	+00125	+00187	
+00007	+00225	+00009	+00119	+00169	
+00011	+00178	+00010	+00089	+00258	
+00006	+00185	+00010	+00118	+00202	
+00006	+00215	+00010	+00141	+00196	
+00003	+00273	+00010	+00153	+00121	
+00011	+00162	+00010	+00075	+00249	
+00008	+00197	+00003	+00092	+00092	
+00005	+00184	+00009	+00160	+00141	
+00002	+00223	+00003	+00109	+00120	
+00001	+00228	+00007	+00127	+00107	
+00010	+00275	+00003	+00182	+00193	
+00007	+00217	+00000	+00124	+00147	
+00004	+00208	+00009	+00089	+00123	
+00000	+00216	+00002	+00096	+00174	
+00009	+00180	+00011	+00123	+00133	

R

79-20-C-2

SAMPLE PROCESSED DATA
EXPERIMENT I

NEITHER FIRST NOR SECOND CHOICE CORRECT FOR INPUT #93

93 7 11 120 7 107 119

NO. UTTERANCES READ= 120 THRESHOLD= 80

WORD#	#OCCUR	ERRORS	REJECT	SCORE	S.D.	DURA	S.D.
0	10	0	0	214.6	12.08	180.6	4.72
	2	6	102.3				
	7	4	106.0				
1	10	0	0	213.6	14.96	111.5	9.85
	4	2	128.5				
	5	6	119.2				
	7	2	128.5				
2	10	0	0	230.1	18.18	119.2	9.24
	0	5	119.2				
	3	2	125.5				
	7	1	120.0				
	8	2	113.0				
3	10	0	1	221.6	62.46	141.9	29.37
	2	2	117.0				
	10	8	150.6				
4	10	0	0	224.7	10.49	143.5	11.17
	1	3	112.7				
	9	7	108.6				
5	10	0	0	195.4	15.21	145.4	9.88
	1	1	137.0				
	9	9	157.6				
6	10	1	0	189.9	24.01	202.7	12.65
	10	9	119.0				
	11	1	120.0				

SECOND CHOICE AVERAGE "CORRELATION"

NUMBER OF OCCURRENCES, SECOND CHOICE

SECOND CHOICE WORD NUMBER

AVERAGE INPUT DURATION

AVERAGE CORRELATION, CORRECT RESPONSES

STANDARD DEVIATION OF "CORRELATIONS"

STANDARD DEVIATION OF DURATIONS

ONLY "ERROR" IN 120 WORDS (WAS ACTUALLY REJECTED BY DURATION TEST)

79-20-C-3

SAMPLE PROCESSED DATA
EXPERIMENT I (Continued)

7	10	0	0	217.5	19.86	161.9	7.15
		0	2	118.5			
		6	6	108.7			
		9	2	125.0			
8	10	0	0	187.6	29.04	85.9	6.09
		2	1	109.0			
		3	8	113.5			
		11	1	106.0			
9	10	0	0	178.0	30.12	141.8	22.23
		4	1	105.0			
		5	2	133.5			
		11	7	104.0			
10	10	0	0	267.0	16.56	209.1	8.41
		3	8	157.4			
		6	2	155.5			
11	10	0	0	168.4	17.64	247.9	27.75
		3	1	46.0			
		8	3	76.3			
		10	6	96.7			

79-20-C-4

PAUSE

APPENDIX D

SAMPLE MESSAGE LIST, EXPERIMENT II

SAMPLE MASTER COMPARISON LISTS, EXPERIMENT II

NARRATIVE MESSAGE STATEMENTS

ENTER DEPARTURE MESSAGE, TRACK 451, TIME 1700, ALTITUDE 350

CHANGE AIRCRAFT IDENTITY OF TRACK 140 TO A27813

ENTER DEPARTURE, TRACK 244 TIME 1509, ALTITUDE 225

HANDOFF TRACK 921 TO SECTOR 66

REQUEST DISPLAY OF FLIGHT PLAN 756

DROP PLAN AND TRACK FOR TRACK 843

AMEND NUMBER 817, IDENTITY ALLEGHENY 0278 TYPE DEHAVILLAND 64,
DISCRETE CODE, TIME 0815 HOURS

ACCEPT HANDOFF, TRACK NUMBER 564

HANDOFF TRACK 445 TO SECTOR 88

ACCEPT HANDOFF, TRACK NUMBER 558

REQUEST STRIP FOR TRACK 377

AMEND COORDINATION FIX AND TIME OF TRACK 310,
LAKE HENRY AT 0547 HOURS

ENTER DEPARTURE MESSAGE, TRACK 448, TIME 0806, ALTITUDE 290

REQUEST WEATHER FOR HAZELTON

REQUEST DISPLAY OF FLIGHT PLAN 939

HOLDING TRACK 359, 0947 HOURS, AT HUGUENOT

AMEND ASSIGNED ALTITUDE OF TRACK 362 TO FLIGHT LEVEL 290

ENTER DEPARTURE MESSAGE, TRACK 756, TIME 0634, ALTITUDE 290

HANDOFF TRACK 632 TO SECTOR 92

REQUEST DISPLAY OF FLIGHT PLAN 713

DROP TRACK AND PLAN FOR TRACK 581

AMEND NUMBER 837, ASSIGNED ALTITUDE 175, SPEED 425,
COORDINATION FIX TOBYHANNA

ACCEPT HANDOFF, TRACK NUMBER 412

HANDOFF TRACK 549 TO SECTOR 90

REQUEST DISPLAY OF FLIGHT PLAN 976

79-20-D-1

MASTER LIST
25 VOICE MESSAGES

	WORDS
DM 451 1700 08 350 (ENTER)	13
AM 140 02 A27813 (ENTER)	12
DM 244 1509 08 225 (ENTER)	13
HO 921 66 (ENTER)	7
FR 756 (ENTER)	5
RS 843 (ENTER)	5
AM 817 02 AL0278 03 DH64 /U 07 0815 (ENTER)	22
HO 564 (ENTER)	5
HO 445 88 (ENTER)	7
HO 558 (ENTER)	5
SR 377 (ENTER)	5
AM 310 06 LHY 07 0547 (ENTER)	12
DM 448 0806 08 290 (ENTER)	13
WR HZL (ENTER)	3
FR 939 (ENTER)	5
HM 359 0947 HUO (ENTER)	10
AM 362 08 290 (ENTER)	9
DM 756 0634 08 290 (ENTER)	13
HO 632 92 (ENTER)	7
FR 713 (ENTER)	5
RS 581 (ENTER)	5
AM 837 08 175 05 425 06 TSD (ENTER)	15
HO 412 (ENTER)	5
HO 549 90 (ENTER)	7
FR 976 (ENTER)	5
R	

79-20-D-2

MASTER LIST
25 KEYBOARD MESSAGES

	CHARACTERS
DM 451 1700 350 (ENTER)	16
AM 140 AID A27813 (ENTER)	18
DM 244 1509 225 (ENTER)	16
HO 921 66 (ENTER)	9
FR 756 (ENTER)	7
RS 843 (ENTER)	7
AM 817 AID AL0278 TYP DH64/U TIM 0815 (ENTER)	38
HO 564 (ENTER)	7
HO 445 88 (ENTER)	9
HO 558 (ENTER)	7
SR 377 (ENTER)	7
AM 310 FIX LHY TIM 0547 (ENTER)	24
DM 448 0806 290 (ENTER)	16
WR HZL (ENTER)	7
FR 939 (ENTER)	7
HM 359 0947 HUO (ENTER)	16
AM 362 ALT 290 (ENTER)	15
DM 756 0634 290 (ENTER)	16
HO 632 92 (ENTER)	9
FR 713 (ENTER)	7
RS 581 (ENTER)	7
AM 837 ALT 175 SPD 425 FIX TSD (ENTER)	31
HO 412 (ENTER)	7
HO 549 90 (ENTER)	9
FR 976 (ENTER)	7
R	79-20-D-3

APPENDIX E

SAMPLES OF RAW DATA, EXPERIMENT II

RAW DATA
25 VOICE MESSAGES

```
DM 451 1700 08 350 (ENTER)                    ERROR (SHOULD BE "8")
+00013
AM 140 02 A27Y13 (ENTER)
+00012
DM *244 1509 08 225 (ENTER)
+00014                                        TIME IN SECONDS, FIRST WORD
HO 921 66 (ENTER)                             TO LAST WORD
+00005
FR 756 (ENTER)
+00004
CN *843 (ENTER)                               REJECT
+00006
AM **817 02 AL0278 03 DH64 */U *07 0*815 (ENTER)
+00052
HO 564 (ENTER)
+00004
HO 445 *88 (ENTER)
+00008
HO 558 (ENTER)
+00005
SR 377 (ENTER)
+00004
AM 310 06 LHY 07 0547 (ENTER)
+00014
DM 448 0806 08 290 (ENTER)
+00012
WR HZL (ENTER)
+00001
**FR 539 ^                                    "ERASE"
FR 939 (ENTER)
+00013
HM 359 **0947 HUO (ENTER)
+00017
AM 362 08 290 (ENTER)
+00008
DM 7*56 063*4 08 290 (ENTER)
+00018
HO 632 92 (ENTER)
+00006
**FR 713 (ENTER)
+00004
CN 5*1_                                       "BACKSPACE"
CN 581 (ENTER)
+00009
AM 837 08 175 05 425 06 TSD (ENTER)
+00021
HO 412 (ENTER)
+00004
HO 549 90 (ENTER)
+00006
*FR 976 (ENTER)
+00005
R
```

79-20-E-1

RAW DATA
25 KEYBOARD MESSAGES

```
DM 5_                                        <- BACKSPACE
DM 451 1700 350 (ENTER)
+00017                                       <- TIME IN SECONDS, FIRST KEY TO LAST KEY
AM 140 AID A27813 (ENTER)
+00016
DM 244 1509 ALT 225 (ENTER)
+00018
HO 921 66 (ENTER)
+00004
FR 756 (ENTER)
+00001
RS 843 (ENTER)
+00003
AM 817 AID AL0278 TYP DH64 /T TIM 0815 (ENTER)   LANGUAGE ERROR WRONG QUALIFIER CODE
+00072                                           FORMAT ERROR (NO SPACE)
HO 564 (ENTER)
+00002
HO 435 88 (ENTER)                            <- KEY ERROR (SHOULD BE "4")
+00004
HO 558 (ENTER)
+00002
RS 377 (ENTER)
+00003
AM 310 FIX LHY TIM 0547 (ENTER)
+00043
DM 448 0806 ALT 29
DM 448 0806 28_
DM 448 0806 290 (ENTER)
+00023
WR HZL (ENTER)
+00010
FR 939 (ENTER)
+00004                                       KEY ERROR (SHOULD BE "4")
HM 359 0937 HUG (ENTER)
+00022
AM 362 ALT 290 (FNIER)
+00010
DM 756 0634 ALT 290 (ENTER)
+00018
HO 632 92 (ENTER)
+00003
FR 713 (ENTER)
+00003
RS 581 (ENTER)
+00005
AM 837 ALT 175 SPD 425 FIX TSD (ENTER)
+00053
HO 412 (ENTER)
+00002
HO 549 90 (ENTER)
+00007
FR 976 (ENTER)
+00003
R
```

79-20-E-2

APPENDIX F

SAMPLES OF PROCESSED DATA, EXPERIMENT II

PROCESSED DATA
25 VOICE MESSAGES

```
INPUT STRING(S)
                                              REJECT
DM *244 1509 08 225 (ENTER)

CN *843 (ENTER)

AM **817 02 AL0278 03 DH64 */U *07 0*815 (ENTER)

HO 445 *88 (ENTER)
                                              "ERASE"
**FR 539 ^
FR 939 (ENTER)
HM 359 **0947 HUO (ENTER)

DM 7*56 063*4 08 290 (ENTER)
**FR 713 (ENTER)
                                              "BACKSPACE"
CN 5*1_
INPUT MSG & MSG TABLE ENTRY DO NOT MATCH
RS 581 (ENTER)
CN 581 (ENTER)                                LANGUAGE ERROR-OPERATOR
DISPLACEMENT:                                 SAID "CANCEL" INSTEAD OF
+00021                                        "DROPTRACK"
*FR 976 (ENTER)

TOTAL BACKSPACE:
+00001
TOTAL ERASE:
+00001
TOTAL ERRORS:
+00001
TOTAL REJECTS:
+00018
TOTAL STRINGS:
+00052
TOTAL CHARACTERS
+00335
TOTAL DURATION
+00250

ERROR SUMMARY:
MASTER CHAR     INPUT CHAR
    8               Y
FUNCTION COMPLETED
TYPE & FOR INSTRUCTIONS
```

79-20-F-1

PROCESSED DATA
25 KEYBOARD MESSAGES

```
INPUT STRING(S)

INPUT MSG & MSG TABLE ENTRY DO NOT MATCH
DM 244 1509 225 (ENTER)
DM 244 1509 ALT 225 (ENTER)
DISPLACEMENT:                        FORMAT ERROR, CODE NOT REQUIRED
+00003

INPUT MSG & MSG TABLE ENTRY DO NOT MATCH
AM 817 AID AL0278 TYP DH64/U TIM 0815 (ENTER)
AM 817 AID AL0278 TYP DH64 /T TIM 0815 (ENTER)
DISPLACEMENT:
+00007                               LANGUAGE ERROR, WRONG CODE

INPUT MSG & MSG TABLE ENTRY DO NOT MATCH
SR 377 (ENTER)                       LANGUAGE ERROR, WRONG CODE
RS 377 (ENTER)
DISPLACEMENT:
+00011

INPUT MSG & MSG TABLE ENTRY DO NOT MATCH
DM 756 0634 290 (ENTER)
DM 756 0634 ALT 290 (ENTER)
DISPLACEMENT:
+00018                               FORMAT ERROR, CODE NOT REQUIRED

TOTAL BACKSPACE:
+00002
TOTAL ERASE:
+00001
TOTAL ERRORS:
+00002
TOTAL REJECTS:
+00000
TOTAL STRINGS:
+00053
TOTAL CHARACTERS
+00267
TOTAL DURATION
+00237

ERROR SUMMARY:
MASTER CHAR     INPUT CHAR
    4               5
    4               3
FUNCTION COMPLETED
TYPE & FOR INSTRUCTIONS
```

79-20-F-2